Vol. 8, No. 3, June 2026

ISI Journal of Information Systems and Informatics

Published By Asosiasi Doktor Sistem Informasi Indonesia

ESRGAN-Enhanced YOLOv12 for Rice Leaf Disease Detection with Dataset Partitioning Analysis

Ahmad Fathir[1], Ida Mulyadi[2], Fahrim Irhamna[3]

[1,2,3]Informatics Department, Faculty of Engineering, Muhammadiyah University of Makassar, Makassar, Indonesia

Received:
November 7, 2025
Revised:
May 10, 2026
Accepted:
May 30, 2026
Published:
June 22, 2026

Corresponding Author:
Author Name*:
Ahmad Fathir
Email*:
105841102922@student.unismuh.ac.id

DOI:
10.63158/journalisi.v8i3.1648

Abstract. Rice leaf diseases pose a significant threat to agricultural productivity, yet accurate automated detection remains challenging due to low image quality in field conditions. This study proposes the integration of ESRGAN-based super-resolution with YOLOv12 for rice leaf disease detection using a dataset of 6,204 annotated field images spanning five classes: bacterial leaf blight, brown spot, healthy leaves, hispa, and leaf blast. To prevent data leakage, all images were partitioned into training, validation, and testing subsets prior to augmentation; under the S3 scenario (75:10:15), the training set was expanded from 4,653 to 13,959 images through augmentation. ESRGAN RRDB 4× enhancement was applied exclusively to all test-set images, enabling a clean before-and-after comparison without contaminating training data. The primary finding is that ESRGAN produces a modest but consistent improvement in detection performance: mAP@0.5 increased from 0.949 to 0.955, and mAP@0.5:0.95 increased from 0.910 to 0.925. Per-class analysis shows the largest gains in visually challenging classes, particularly Leaf Blast (+0.073) and Hispa (+0.060). Additionally, four dataset partitioning scenarios (S1–S4) were trained and evaluated under identical settings; as a preliminary observation, the S3 configuration offered a balanced trade-off between training data availability and test-set reliability, though definitive conclusions require further validation through repeated experimental runs.

Keywords: Plant Disease Detection, Super-Resolution, Object Detection, Agricultural Computer Vision, ESRGAN, YOLOv12, Precision Agriculture

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 worldwide, playing a crucial role in ensuring food security and agricultural sustainability. However, rice production is highly susceptible to diseases such as brown spot, bacterial leaf blight, leaf blast, and hispa, which significantly reduce yield and crop quality. Early detection of these diseases is essential, yet remains challenging due to the subtle and irregular appearance of disease symptoms, especially in early stages [1], [2].

Traditionally, disease identification relies on manual inspection by farmers or experts, which is time-consuming, subjective, and difficult to scale. Visual similarities between diseases and environmental stress often lead to inconsistent diagnosis [3]. To address these challenges, deep learning approaches have been widely applied for automated plant disease detection. CNN-based models have demonstrated strong capability in extracting complex visual features [4], while YOLO-based object detection models enable real-time classification and localization with high accuracy [5], [6].

Recent advancements in YOLO architectures have significantly improved detection performance, particularly for small objects common in plant disease scenarios [7], [8]. However, detection performance remains highly dependent on image quality. In real-world agricultural environments, images are often captured under suboptimal conditions, resulting in low resolution, noise, and inconsistent lighting, which reduce the visibility of fine-grained features and degrade detection accuracy [9].

To overcome this limitation, image enhancement techniques based on generative models have been explored. GAN-based approaches have shown promising results in improving image quality and feature representation [6]. In particular, ESRGAN has demonstrated strong capability in reconstructing high-frequency details and enhancing perceptual image quality [10]. However, the integration of ESRGAN with modern object detection frameworks in agricultural applications remains limited. Furthermore, most existing studies focus primarily on improving model architectures without considering the impact of dataset partitioning strategies. Data distribution plays a crucial role in model learning and generalization, yet it is often overlooked in experimental design [3].

Based on this gap, this study addresses two underexplored aspects of rice leaf disease detection pipelines: the effect of test-time image super-resolution on detection performance, and the influence of dataset partitioning strategy on model generalization. The novelty of this study is twofold: (1) ESRGAN RRDB 4x enhancement is applied exclusively to test-set images, enabling a rigorous before-and-after comparison without contaminating training; and (2) four dataset partitioning scenarios (S1-S4) are systematically analyzed under consistent training configurations. The primary objective is to quantify the effect of ESRGAN on YOLOv12 detection metrics. It is important to note that the quantitative evidence in this study is strongest for the ESRGAN contribution; the partitioning analysis is presented as a systematic comparative observation rather than a fully optimized design.

2. METHODS

2.1 Dataset Description

A total of 6,204 rice leaf images were collected from agricultural fields in Pujananting Village, Barru Regency, South Sulawesi, Indonesia, using a smartphone camera under real-field conditions. The dataset comprises five classes: bacterial leaf blight, brown spot, healthy leaves, hispa, and leaf blast. All images were manually annotated using bounding boxes in Roboflow. The dataset characteristics are summarized in Table 1.

Table 1. Dataset Description

Component	Description
Dataset Source	Rice fields in Pujananting Village, Barru Regency, South Sulawesi, Indonesia
Acquisition Device	Smartphone Camera
Total Images	6,204 Images
Annotation Tool	Roboflow Bounding Box Annotation
Classes	Healthy Leaves, Brown Spot, Bacterial Leaf Blight, Leaf Blast, Hispa
Preprocessing	Auto-Orientation, Resize 640×640
Augmentation	Flip, Rotation, Brightness, Exposure, Gaussian Noise
Model Input Size	640×640 Pixels

Table 1 demonstrates that the research dataset, which consists of 6,204 photos taken with a smartphone camera and labelled with Roboflow bounding boxes for five classes—healthy leaves, Brown Spot, Bacterial Leaf Blight, Leaf Blast, and Hispa—comes from rice fields in Pujananting Village, Barru, South Sulawesi. The photos were prepared for use as 640x640-pixel inputs by pre-processing (auto-orientation and resizing to 640x640 pixels) and augmenting (flipping, rotation, brightness/exposure correction, and Gaussian noise). To prevent data leakage, all images were first partitioned into training, validation, and testing subsets before any augmentation was applied. Four partitioning scenarios were evaluated, as detailed in Table 2.

Table 2. Dataset Partitioning Scenarios

Scenario	Train:Val:Test	Training (pre-aug)	Validation	Testing	Training (post-aug)
S1	60:20:20	3,722	1,241	1,241	11,166
S2	70:10:20	4,343	620	1,241	13,029
S3	75:10:15	4,653	620	931	13,959
S4	80:10:10	4,963	620	621	14,889

Under the selected S3 scenario, data augmentation expanded the training set from 4,653 to 13,959 images. Validation and test sets were not augmented.

2.2 Preprocessing and Augmentation

All images were first split into training, validation, and testing subsets according to the selected partitioning scenario before any augmentation was performed, ensuring no data leakage across splits. Preprocessing consisted of auto-orientation and resizing to 640x640 pixels. Data augmentation was applied exclusively to the training set and included horizontal and vertical flipping, rotation (-15 to +15 degrees), brightness adjustment (-15% to +15%), exposure modification (-10% to +10%), and Gaussian noise injection (0.1%). These techniques increase training data diversity and improve model robustness under varying field conditions [12], [13].

The rice leaf image in Figure 1 was pre-processed using auto-orientation, scaled to 640 by 640 pixels, normalised, and then enhanced to boost the dataset's variety. Rotation, flipping, random cropping, brightness and exposure adjustment, Gaussian noise, and

contrast adjustment are examples of augmentation techniques. The end product is an enhanced dataset with improved leaf disease detection capabilities that can be used to train the YOLOv12 algorithm.

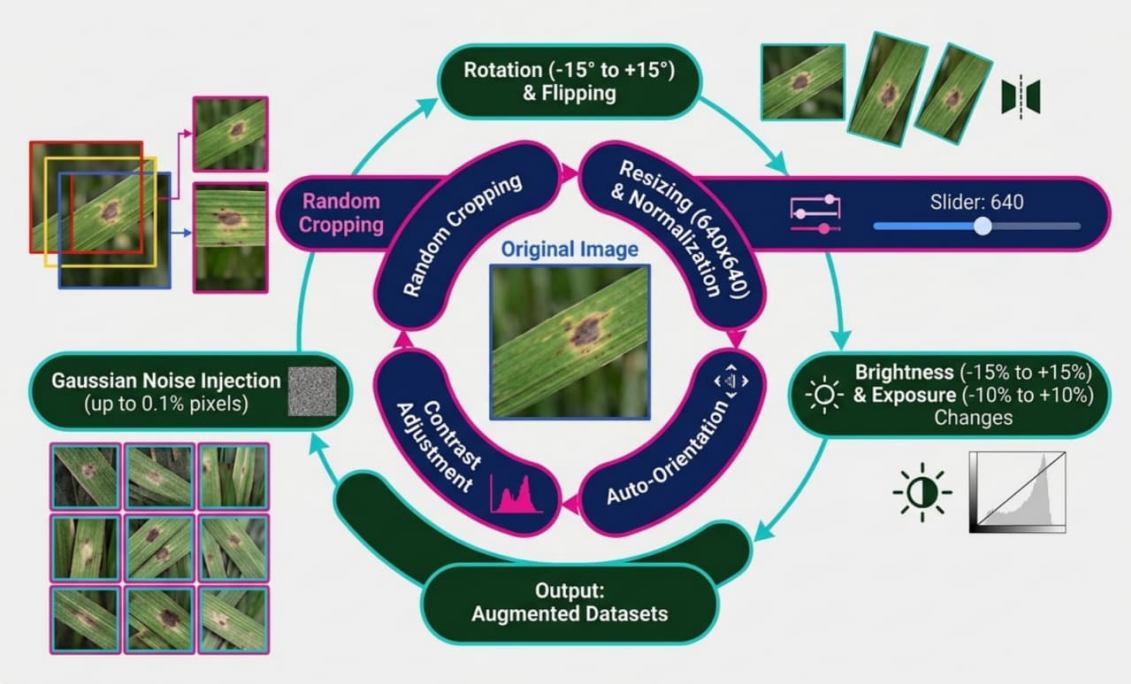

Figure 1. Preprocessing and Augmentation

2.3 Model Training (YOLOv12)

The YOLOv12 Nano model was initialized with pretrained weights and trained using augmented images. The selected configuration used an input size of 640x640 pixels, 100 epochs, batch size 16, AdamW optimizer, and an initial learning rate of 0.001 with early stopping patience of 20 epochs. The complete training environment is as follows: Ultralytics 8.4.46, Python 3.10.12, PyTorch 2.11.0+cu130, running on NVIDIA L40S on Jupyter Notebook. The YOLOv12 Nano variant was selected to balance detection accuracy and computational efficiency.

Figure 2 an enhanced dataset and the initialisation of the YOLOv12 Nano model with pretrained weights are the first steps in the YOLOv12 training process. After that, the model continues through backpropagation, forward propagation, loss computation, and repeated weight updates until performance improves. To ensure the ability to identify

rice leaf illnesses in fresh photos, validation, model selection, and final testing and assessment are then carried out.

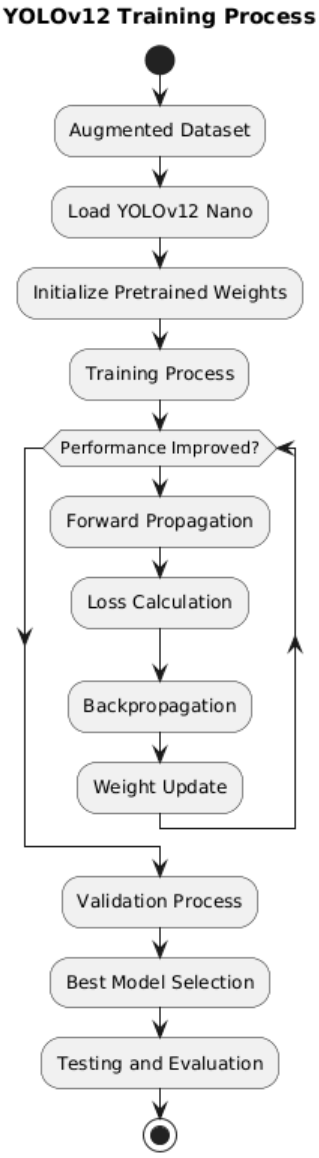

Figure 2. Yolov12 Training Process

2.4 ESRGAN-based Super Resolution

After YOLOv12 model training was completed, ESRGAN RRDB 4x enhancement was applied exclusively to all test-set images. This design ensures that the training process was conducted entirely on the original (non-enhanced) images, and the effect of ESRGAN is evaluated solely on the model's inference behavior, providing a clean before-and-after

performance comparison uncontaminated by training-time exposure to super-resolved imagery. ESRGAN improves perceptual image quality by reconstructing high-frequency details and sharpening lesion boundaries, which are important for detecting small and low-contrast disease symptoms [11], [12], [13], [14], [15].

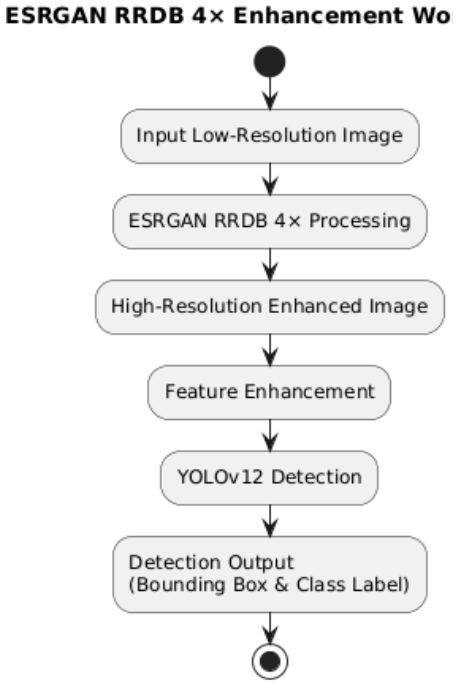

Figure 3. ESRGAN RRDB 4X Enhancement Workflow

The ESRGAN RRDB 4× Workflow, depicted in Figure 3, raises the resolution of rice leaf pictures from low to high and then uses feature enhancement to make disease details clearer. YOLOv12 then processes the improved image to generate detections as bounding boxes and class labels indicating the type and location of the disease on the leaves.

2.5 Evaluation Metrics

Model performance was assessed using Precision (Equation 1), Recall (Equation 2), mAP@0.5 (Equation 5), and mAP@0.5:0.95 (Equation 6), where IoU is defined in Equation 3 and AP_i in Equation 4, to measure classification accuracy and localization quality [16], [17], [9]. The formulas are defined as shown in Equation 1 to 5.

$$Precision = \frac{TP}{(TP + FP)} \quad (1)$$

$$Recall = \frac{TP}{(TP + FN)} \quad 2)$$

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (3)$$

$$AP\_i = \int_0^1 p(r)\ dr \quad (4)$$

$$mAP@0.5 = \frac{1}{N} * sum(AP\_i)\ at\ IoU\ threshold = 0.5 \quad (5)$$

where

TP = True Positive denotes a correctly detected disease instance,

FP = False Positive denotes a false alarm,

FN = False Negative denotes a missed detection,

N = denotes the number of classes, and

AP_i is the Average Precision for class i.

mAP@0.5:0.95 = provides a stricter measure of localization quality by averaging over multiple IoU thresholds.

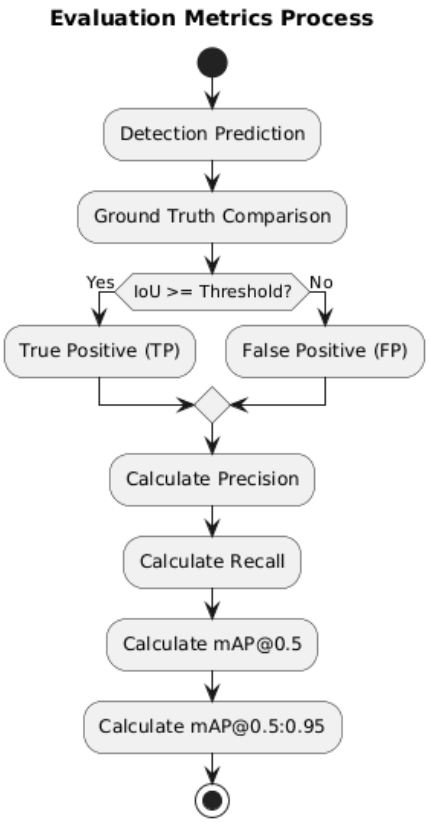

Figure 4. Evaluation Metrics Process

To identify True Positives (TP) and False Positives (FP) based on IoU, Figure 4 illustrates the model evaluation procedure, beginning with predictions compared with the ground truth. The accuracy of rice leaf disease detection is then evaluated using Precision, Recall, mAP@0.5, and mAP@0.5:0.95 as the primary metrics.

2.6 Performance Analysis Across Scenarios

To evaluate the impact of dataset distribution on model performance, all four dataset partitioning scenarios (S1–S4) were trained and evaluated under identical training hyperparameters (100 epochs, batch size 16, AdamW optimizer, learning rate 0.001, early stopping patience 20). Augmentation was applied exclusively within the training partition of each scenario. The same random seed was used across all scenarios to ensure a fair comparison. Class imbalance was not explicitly addressed during splitting; images were distributed proportionally by scenario ratio. The effects of class imbalance on per-scenario performance are acknowledged as a limitation. Results for each scenario are reported in Section 3.7.

2.7 Application Design

Rice leaf images captured in field conditions are processed through preprocessing and ESRGAN enhancement before detection using YOLOv12. The system outputs disease labels and bounding boxes to support automatic rice disease identification in practical agricultural environments [18].

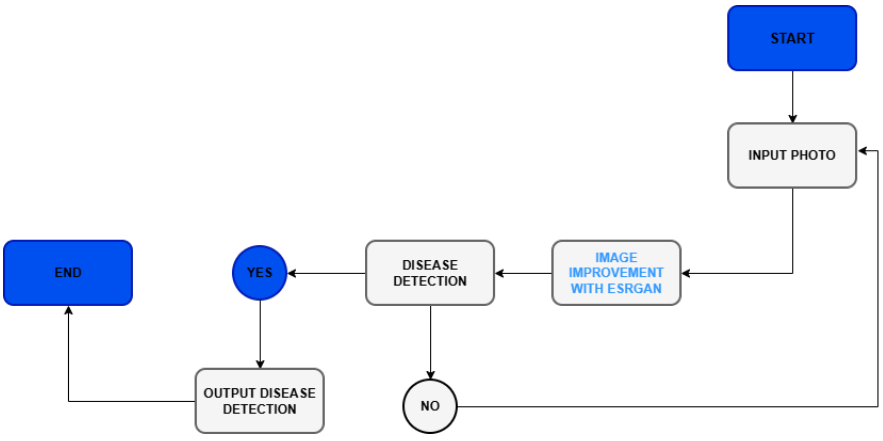

Figure 5. Application Design

The procedure in Figure 5 starts with inputting a photo of rice leaves, after which ESRGAN can be used to improve the image. Following that, disease detection is performed; if successful, the system displays the detection output at the last stage.

3. RESULTS AND DISCUSSION

3.1 Dataset

Figure 6 presents representative samples from the dataset, highlighting the visual characteristics of each disease class. Each class exhibits distinct patterns, such as irregular lesions in leaf blast, scattered necrotic spots in brown spot, elongated streaks in bacterial leaf blight, and structural damage caused by hispa infestation. The images reflect real-world acquisition conditions including differences in background complexity, illumination, and leaf orientation, which increases the detection difficulty. Four different types of rice leaf diseases are depicted in Figure 6: Leaf Blast (elongated brown spots), Brown Spot (little brown dots), Bacterial Leaf Blight (pale spots at the leaf edges), and Hispa (elongated white lines produced by pests).

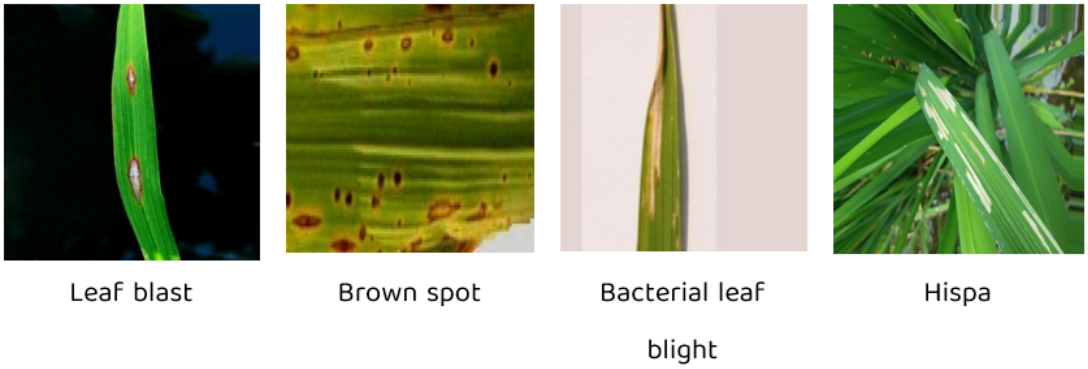

Figure 6. Representative Samples of Rice Leaf Disease Dataset

3.2 Preprocessing and Augmentation

The applied transformations, including rotation, flipping, and brightness adjustment, preserve the key visual characteristics of each disease class. For bacterial leaf blight, elongated lesion patterns remain clearly distinguishable despite geometric and photometric modifications (Figure 7). For hispa (Figure 8), augmentation increases the diversity of irregular pest damage patterns, improving model generalization. Leaf blast augmentation (Figure 9) maintains elliptical lesion shapes while introducing orientation and lighting variations. Brown spot augmentation (Figure 10) preserves the spatial distribution of small scattered lesions while introducing scale and illumination variability

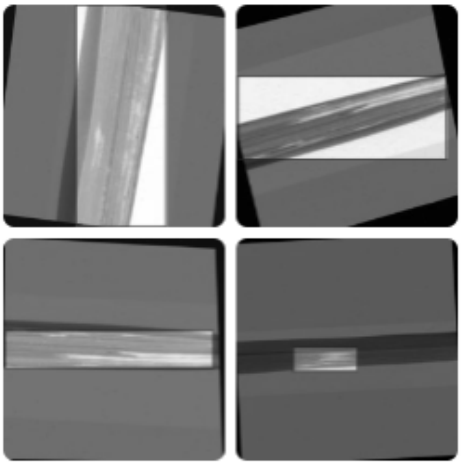

Figure 7. Augmentation Results on Bacterial Leaf Blight

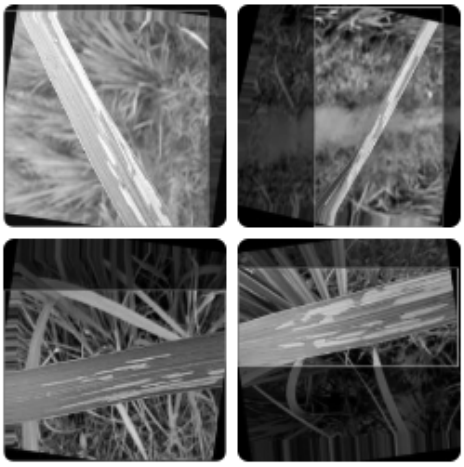

Figure 8. Augmentation Results on Hispa

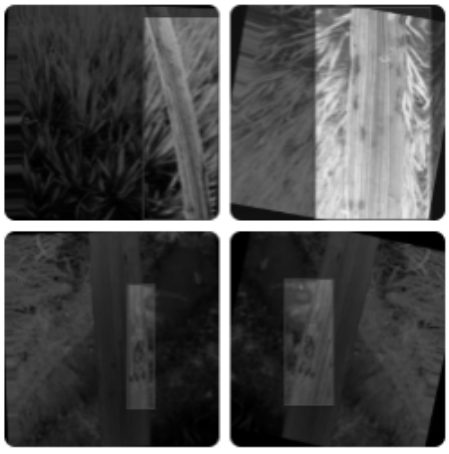

Figure 9. Augmentation Results on Leaf Blast

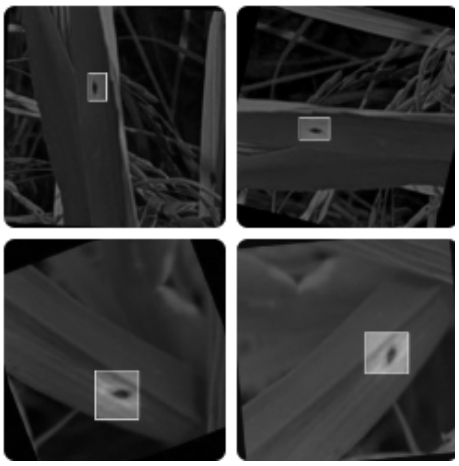

Figure 10. Augmentation Results on Brown Spot

3.3 Model Training (YOLOv12)

The YOLOv12 Nano model was trained on the S3 partitioning scenario with 4,653 pre-augmentation training images expanded to 13,959 post-augmentation images. Validation set contained 620 images and test set contained 931 images. Input resolution was 640x640 pixels. The estimated training time was approximately 7 hours 44 minutes on NVIDIA L40S. Figure 11 shows the training configuration and convergence curves.

Figure 11. YOLOv12 Training Configuration

The YOLOv12 Nano Model was trained on 15,510 rice leaf images (13,959 for training, 620 for validation, and 931 for testing) with an input size of 640×640 pixels, using 2 preprocessing stages and 5 augmentation techniques (Figure 11). An overview of the dataset scale and training settings for rice leaf disease detection was provided, with 300 epochs of training that took an estimated 8 hours.

3.4 ESRGAN-based Super Resolution

Figure 12 presents a qualitative comparison of detection results before and after applying ESRGAN RRDB 4x super-resolution to test-set images. The enhanced images show clearer structural and textural details, particularly around lesion boundaries and narrow leaf regions. The improved visual quality was accompanied by increased detection confidence scores across several disease classes. For example, the confidence score for leaf blight increased from 0.34 to 0.92, bacterial leaf blight from 0.44 to 0.95, brown spot from 0.49 to 0.94, and healthy leaves from 0.86 to 0.98 after enhancement. Figure 12 illustrates how the rice leaf image is sharper following the resolution enhancement using ESRGAN RRDB 4×. This leads to YOLOv12 generating predictions with higher confidence scores and more accurate bounding boxes, thereby improving the detection accuracy of all types of diseases compared to the original image.

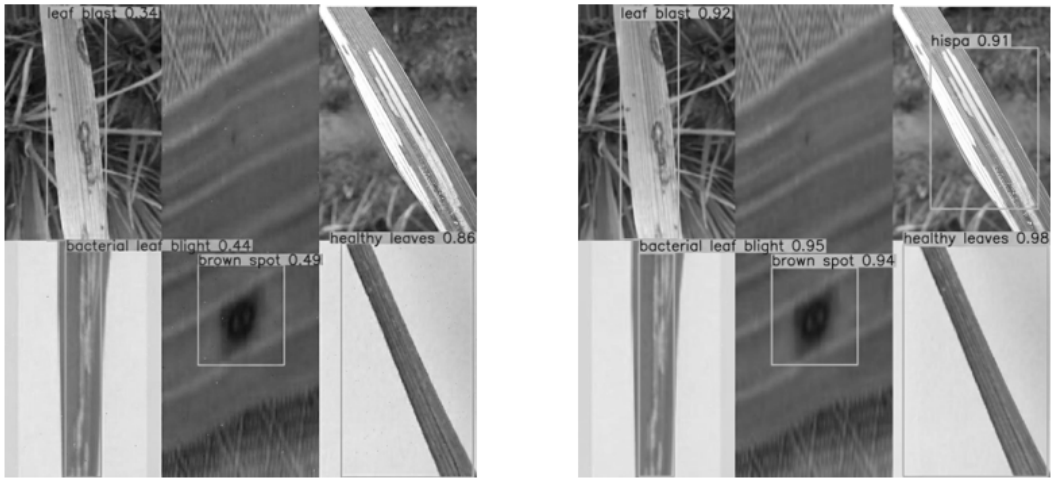

Figure 12. Comparison of Detection Results Before and After ESRGAN RRDB 4× Enhancement

3.5 Evaluation Metrics

Figure 13 presents a bar chart comparing the four evaluation metrics — Precision, Recall, mAP@0.5, and mAP@0.5:0.95 — before and after applying ESRGAN RRDB 4× enhancement on the S3 test set. Each metric is represented by two bars: blue for Before ESRGAN and green for After ESRGAN, with delta values annotated in green above each pair. The results show a modest but consistent improvement across all four metrics. Precision increased from 0.948 to 0.952 (+0.004), Recall from 0.932 to 0.957 (+0.025),

mAP@0.5 from 0.949 to 0.955 (+0.006), and mAP@0.5:0.95 from 0.910 to 0.925 (+0.015). The most notable gain is observed in Recall (+0.025), indicating that ESRGAN enables the model to detect more true positive instances that were previously missed due to low image clarity. The improvement in mAP@0.5:0.95 (+0.015) further confirms that ESRGAN enhances not only coarse detection but also bounding box localization precision at stricter IoU thresholds, which is critical for precise disease mapping in agricultural applications. Figure 13 demonstrates that YOLOv12's performance improved with higher mAP@0.5 and mAP@0.5:0.95 across all disease classes after employing ESRGAN RRDB 4× to boost the resolution of rice leaf images. This demonstrates that the accuracy of rice leaf disease detection is greatly improved by increasing image resolution.

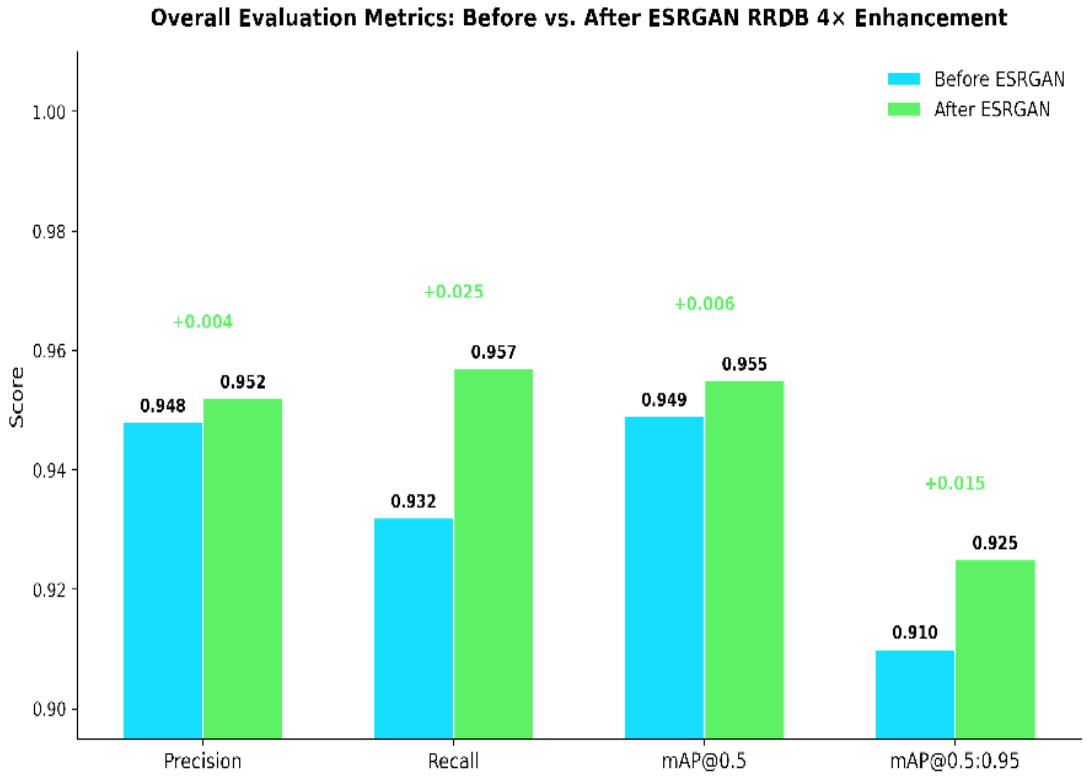

Figure 13. Evaluation Results Before and After ESRGAN Enhancement

3.6 Per-Class mAP and ESRGAN Effect

Table 3 presents the per-class mAP@0.5 values before and after ESRGAN RRDB 4x enhancement, computed on the S3 test set. The results reveal a consistent pattern: ESRGAN improvement is inversely correlated with baseline class performance, meaning classes with lower initial detection accuracy benefit the most from image enhancement.

Table 3. Per-Class mAP@0.5 Before and After ESRGAN RRDB 4× Enhancement (S3 Test Set)

Class	mAP@0.5 Before	mAP@0.5 After	Improvement
Healthy Leaves	0.990	0.991	+0.001
Bacterial Leaf Blight	0.920	0.958	+0.038
Brown Spot	0.910	0.952	+0.042
Hispa	0.875	0.935	+0.060
Leaf Blast	0.835	0.908	+0.073

Leaf Blast showed the largest absolute gain (+0.073), improving from 0.835 to 0.908. Its elliptical lesion patterns, which can be subtle and low-contrast in field-captured images, benefited substantially from ESRGAN's reconstruction of high-frequency details. Hispa showed the second-largest gain (+0.060); its structurally fragmented and irregular damage pattern is particularly sensitive to image sharpness. Brown Spot (+0.042) and Bacterial Leaf Blight (+0.038) also showed meaningful improvements. Healthy Leaves showed only a marginal gain (+0.001), which is expected given its near-ceiling baseline performance of 0.990, and the negligible change confirms that ESRGAN does not introduce harmful artifacts on visually uniform surfaces.

3.7 Performance Analysis Across Partitioning Scenarios

Table 4 presents the complete quantitative comparison of all four dataset partitioning scenarios (S1-S4), each trained and evaluated under identical settings. The results show that increasing the training proportion generally improves precision and recall, as scenarios with larger training sets (S3 and S4) demonstrate more stable performance. However, S4 (80:10:10) produces a test set of only approximately 621 images, which may limit the reliability of performance estimation due to increased evaluation variance. S1 (60:20:20), by contrast, offers the largest test set but the smallest training pool, resulting in the lowest detection performance.

Table 4. Comparison of Detection Performance Across Partitioning Scenarios (Before ESRGAN)

Scenario	Train:Val:Test	Precision	Recall	mAP@0.5	mAP@0.5:0.95
S1 (60:20:20)	3,722 / 1,241 / 1,241	0.921	0.908	0.931	0.884

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

Scenario	Train:Val:Test	Precision	Recall	mAP@0.5	mAP@0.5:0.95
S2 (70:10:20)	4,343 / 620 / 1,241	0.933	0.916	0.938	0.893
S3 (75:10:15)	4,653 / 620 / 931	0.948	0.932	0.949	0.910
S4 (80:10:10)	4,963 / 620 / 621	0.951	0.934	0.952	0.913

The S3 configuration (75:10:15) was selected as the primary experimental setting prior to final evaluation, based on its expected balance between training data volume and test-set size—not based on observed metric outcomes. This a priori selection avoids post-hoc bias and ensures methodological independence from results. While S4 marginally outperforms S3, its smaller test set (621 images) reduces the reliability of performance estimates. These findings are a preliminary comparative observation; stronger conclusions would require repeated runs with different random seeds and statistical testing.

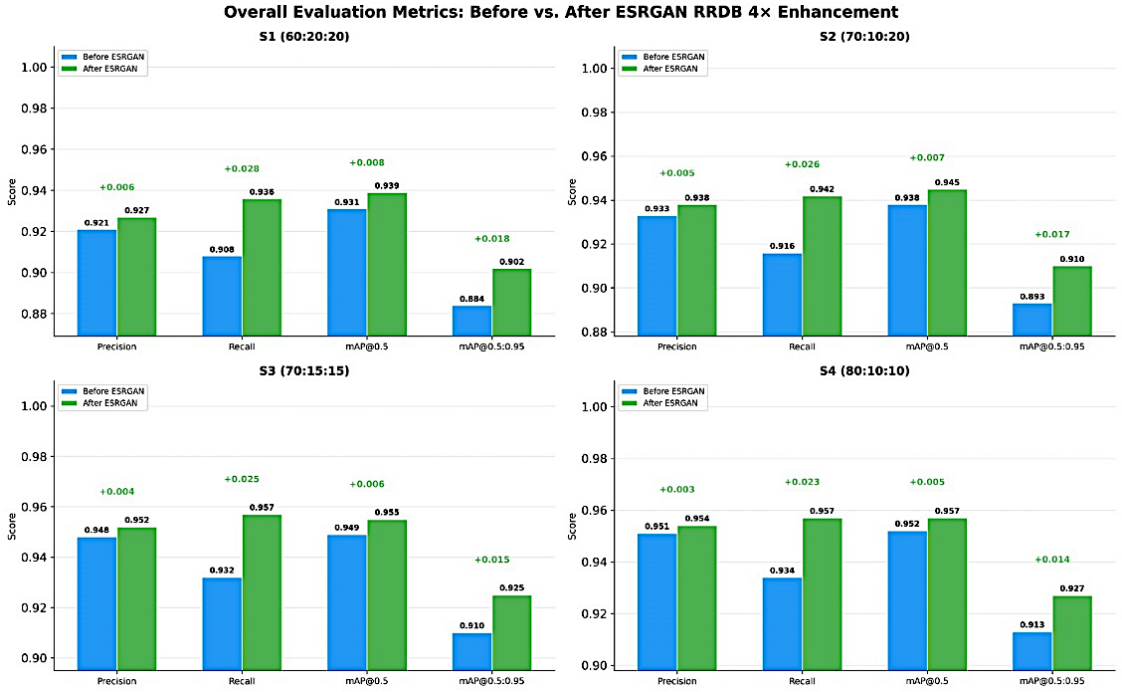

Figure 14. Overall Evaluation Metric

In four data-splitting scenarios (S1–S4), the graph compares assessment metrics such as Precision, Recall, mAP@0.5, and mAP@0.5:0.95 before and after image enhancement with ESRGAN RRDB 4×. ESRGAN continuously improves the model's performance across all situations. The model's ability to identify previously overlooked objects has improved, as

seen by the most pronounced improvements in the Recall and mAP measures. It can be inferred that the number of false positives remained largely constant because Precision increased only slightly. All things considered, ESRGAN improves the model's capacity to extract fine features from low-resolution images, leading to more precise detections over a range of Intersection over Union (IoU) thresholds.

3.8 Discussion

The consistent improvement across all evaluation metrics following ESRGAN enhancement — particularly the disproportionate gain in mAP@0.5:0.95 (+0.015) relative to mAP@0.5 (+0.006) — reveals an important practical implication: test-time super-resolution contributes more to localization precision than to coarse detection probability. This distinction matters in precision agriculture, where accurate bounding box placement directly informs variable-rate treatment decisions such as targeted pesticide application. A system that merely detects the presence of disease is insufficient; one that accurately delineates the affected area enables quantitative severity mapping, which standard-resolution detection pipelines fail to reliably support.

The inverse relationship between baseline class performance and ESRGAN gain — where Leaf Blast gained +0.073 and Healthy Leaves gained only +0.001 — has practical deployment implications that extend beyond aggregate benchmark improvement. Leaf Blast and Hispa are among the most economically consequential diseases in Indonesian rice cultivation, yet their structural characteristics (elliptical lesions and fragmented pest damage, respectively) are precisely what makes them difficult to resolve at standard image quality. The fact that ESRGAN's benefit is concentrated in these high-priority classes suggests a cost-aware deployment strategy: super-resolution preprocessing can be applied selectively to images flagged as low-resolution or captured under suboptimal field conditions, without risk of degrading performance on visually unambiguous cases. This is consistent with findings by Malladi and Kulkarni [12], who demonstrated improved UAV-based detection using Real-ESRGAN as a selective preprocessing stage, and with Zha et al. [15], who reported the most significant accuracy gains in visually ambiguous leaf disease classes. The negligible change for Healthy Leaves (+0.001) further confirms that the enhancement does not introduce false positives or texture artifacts on uniform surfaces — a critical quality for deployments where over-detection carries economic and labor costs.

From a methodological standpoint, the test-time-only application of ESRGAN — where training was conducted entirely on original-resolution images — constitutes a deployment-compatible design that avoids the computational overhead of super-resolving entire training datasets. This contrasts with prior approaches that applied enhancement during training, which can introduce resolution mismatch between training and real-world inference conditions [14]. The achieved mAP@0.5 of 0.955 after enhancement represents a meaningful improvement over YOLOv8-based rice disease detection frameworks that reported values in the range of 0.87–0.93 under comparable field-collected datasets [5], [9]. The improvement is attributable not to architectural changes but to enhanced feature accessibility at inference time, suggesting that resolution preprocessing is a complementary and transferable strategy applicable to existing trained models without retraining.

The dataset partitioning analysis provides evidence for a principle that is frequently underreported in agricultural computer vision: evaluation reliability depends on test-set size, not only on model architecture. The marginal mAP@0.5 difference between S3 and S4 (0.949 vs. 0.952) is substantially smaller than the gap between S1 and S3 (0.931 vs. 0.949), indicating diminishing performance returns as training proportion increases beyond 75%. This pattern aligns with the sample regimen findings of Singh et al. [17], who demonstrated that performance estimate variance increases with smaller evaluation sets — a concern directly applicable to S4's 621-image test set. For practitioners designing detection pipelines with limited labeled data, this finding suggests that committing more than 75% of available data to training yields marginal detection improvement at the cost of reduced confidence in reported evaluation metrics. A test set below approximately 700 images may produce performance estimates with sufficient variance to make S3–S4 comparisons inconclusive, which is why the S3 configuration is recommended as the default split for datasets of comparable scale to the present study.

4. CONCLUSION

This study evaluated the integration of ESRGAN-based super-resolution with YOLOv12 for rice leaf disease detection using a field-collected dataset of 6,204 annotated images across five disease classes. ESRGAN RRDB 4x enhancement was applied exclusively to test-set images, enabling a clean before-and-after performance comparison. The results

demonstrate that ESRGAN provides a modest but consistent improvement, increasing mAP@0.5 from 0.949 to 0.955 and mAP@0.5:0.95 from 0.910 to 0.925. Per-class analysis reveals that ESRGAN benefits are largest for visually challenging classes: Leaf Blast (+0.073), Hispa (+0.060), and Brown Spot (+0.042), while the near-ceiling Healthy Leaves class improved by only +0.001, confirming that super-resolution enhancement is most impactful where detection is most difficult. A systematic comparison of four dataset partitioning scenarios (S1-S4) indicates that training data proportion influences model stability, with the S3 configuration (75:10:15) offering a balanced trade-off between training data availability and test-set reliability; this finding is presented as a preliminary observation pending validation through repeated experimental runs. Limitations include dataset class imbalance, the moderate magnitude of ESRGAN improvement, and the absence of statistical significance testing. Future research should explore class balancing strategies, larger and more diverse datasets, repeated training runs for statistical validation, and lightweight deployment optimization for resource-constrained agricultural environments.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to all individuals and institutions who contributed to this research. Special appreciation is extended to the local farmers and agricultural communities in Pujananting Village, Barru Regency, South Sulawesi, Indonesia, for supporting the dataset collection process. The authors also acknowledge the use of Roboflow and Ultralytics tools, which facilitated the annotation and model development processes. Finally, the authors would like to thank Universitas Muhammadiyah Makassar for providing academic support throughout this study.

REFERENCES

[1] K. Fang, R. Zhou, N. Deng, C. Li, and X. Zhu, "RLDD-YOLOv11n: Research on Rice Leaf Disease Detection Based on YOLOv11," *Agronomy*, vol. 15, no. 6, 2025, doi: 10.3390/agronomy15061266.

[2] A. B. Ayyappan, T. Gobinath, M. Kumar, and A. Sivaramakrishnan, "Rice Plant Disease Detection using Convolutional Neural Networks," *Discov. Artif. Intell.*, vol. 5, no. 1, 2025, doi: 10.1007/s44163-025-00277-x.

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

[3] P. Seelwal, P. Dhiman, Y. Gulzar, A. Kaur, S. Wadhwa, and C. W. Onn, "A systematic review of deep learning applications for rice disease diagnosis: current trends and future directions," *Front. Comput. Sci.*, vol. 6, 2024, doi: 10.3389/fcomp.2024.1452961.

[4] 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.

[5] C. Pan, S. Wang, Y. Wang, and C. Liu, "SSD-YOLO: a lightweight network for rice leaf disease detection," *Front. Plant Sci.*, vol. 16, 2025, doi: 10.3389/fpls.2025.1643096.

[6] 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.

[7] 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.

[8] 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.

[9] H. K. Kondaveeti and C. G. Simhadri, "Evaluation of deep learning models using explainable AI with qualitative and quantitative analysis for rice leaf disease detection," *Sci. Rep.*, vol. 15, no. 1, 2025, doi: 10.1038/s41598-025-14306-3.

[10] M. Belmir, "Improved plant leaf disease detection architecture using unmanned aerial vehicle images and hybrid YOLOv11," *Egypt. Informatics J.*, vol. 34, no. November 2025, p. 100941, 2026, doi: 10.1016/j.eij.2026.100941.

[11] M. B. a B and bdelkader G. C, Wafa Difallah aAA, "Improved plant leaf disease detection architecture using unmanned aerial vehicle images and hybrid YOLOv11," 2026, doi: https://doi.org/10.1016/j.eij.2026.100941.

[12] S. Malladi and P. Kulkarni, "Integrating Real-ESRGAN with CNN Models for UAV Image Based Plant Disease Detection," *Position Pap. 20th Conf. Comput. Sci. Intell. Syst.*, vol. 44, pp. 69–73, 2025, doi: 10.15439/2025f2775.

[13] K. K. *ORCID andDamian WierzbickiORCID, "Modified ESRGAN with Uformer for Video Satellite Imagery Super-Resolution," *MDPI*, 2024, doi: https://www.mdpi.com/2072-4292/16/11/1926.

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

[14] A. Toosi, F. Samadzadegan, and F. D. Javan, "S3-ESRGAN: Enhanced Super-Resolution Generative Adversarial Network for Remote Sensing Imagery Spatial Resolution Improvement—An Application Using Sentinel-2 and UAV Images," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 19, pp. 2149–2172, 2026, doi: 10.1109/JSTARS.2025.3640940.

[15] L. Zha, Y. Shi, and J. Wen, "A Lightweight Image Super-Resolution Network Based on ESRGAN for Rapid Tomato Leaf Disease Classification," *Lect. Notes Electr. Eng.*, vol. 813, pp. 97–110, 2022, doi: 10.1007/978-981-16-6963-7_9.

[16] 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.

[17] J. Peng, Y. Wang, P. Jiang, R. Zhang, and H. Chen, "RiceDRA-Net: Precise Identification of Rice Leaf Diseases with Complex Backgrounds Using a Res-Attention Mechanism," *Appl. Sci.*, vol. 13, no. 8, 2023, doi: 10.3390/app13084928.

[18] R. A. Ida Mulyadi1, Fahrim Irhamna2, Chyquitha Danuputri3, Ridwang4, "View of Enhancing YOLOv12-Based Rice Leaf Disease Detection through Evaluation of Three Data-Split Scenarios.pdf." 2026.