Academic Journal of Science and Technology
ISSN: 2771-3032 | Vol. 21, No. 1, 2026

An Automated Annotation and Lightweight Detection Framework for Rice Leaf Disease Based on Color Space Analysis and YOLOv11

Ningjiang Han, Jiahui Jiang, Zhengzhe Jia, Haodong Yin, Junjie Yang

Chengdu Jincheng College, Chengdu 617737, China

Abstract: Accurate and timely detection of rice leaf diseases is essential for precision agriculture and food security. However, existing object detection approaches heavily rely on labor-intensive manual bounding box annotations, which limits their scalability and practical deployment. In this paper, we propose an automated annotation and detection framework that combines color space analysis-based lesion localization with the lightweight YOLOv11n model for rice leaf disease detection. Our automated annotation method leverages disease-specific HSV color space characteristics and morphological processing to generate bounding box labels without human intervention, achieving a 98.6% successful annotation rate across 5,932 images. The annotated dataset is then used to train YOLOv11n, which achieves a mean Average Precision (mAP@0.5) of 93.36% and mAP@0.5:0.95 of 77.53% on the test set, with only 2.58M parameters and 1.0ms inference time per image. Experimental results on four rice disease categories (Bacterial Blight, Blast, Brown Spot, andTungro) demonstrate that the proposed framework provides an effective and efficient solution for rice disease detection with minimal annotation cost, making it suitable for real-time deployment on edge devices in agricultural settings.

Keywords: Disease detection of rice leaves; automatic labeling; YOLOv11; color space analysis; precision farming; an object detector with low weights.

1. Introduction

Rice has become one of the world most valued staple foods, with more than half of the world population being fed by it. Nevertheless, rice cultivation is exposed to considerable risks due to different diseases such as Bacterial Blight, Rice Blast, Brown Spot, and Tungro, which may lead to yield losses of up to 10-50 percent unless they are identified and addressed in good time [1].The conventional method of disease diagnosis involves experienced agronomists carrying out visual examinations in the field, which is time-consuming, biased, and cannot be applied to large scale farms.

In the last few years, deep learning-based methods of object detection, especially the YOLO (You Only Look Once) series, have demonstrated promising results in the detection of agricultural diseases [2,4]. The models have been used with success to identify and place disease lesions on leaves of crops. Proposed by Jia et al[5], MobileNet-CA-YOLO is an enhanced version of YOLOv7 that incorporates the MobileNetV3 backbone and Coordinate Attention mechanism to detect rice pests and diseases, with a 93.7 per cent mAP@0.5. Nonetheless,

these strategies have a similar bottleneck: both need massive human efforts and domain knowledge to generate large-scale datasets comprising annotated bounding boxes with high precision. The annotation bottleneck is especially severe in agricultural use because of various factors, such as: (1) the disease lesions are usually irregular shaped and small sized, scattered over the surface of the leaves; (2) expert diagnosis is necessary to differentiate between the disease signs and normal leaf variations; and (3) the variety of disease symptoms at different growth stages and in different environments complicates the uniform annotation process.

In order to overcome this issue, our proposal is a new system that combines automated color space-based lesion annotation with the latest version of the YOLOv11n model. Our main contributions include:

An automated disease lesion annotation algorithm that uses disease specific HSV color space analysis and morphological processing, that produces a bounding box annotation without human intervention, with 98.6 percent coverage.

A light weight detection system based on YOLOv11n with 93.36 percent mAP@0.5 (with just 2.58M parameters) and 5.21 MB model size, allowing real-time inference on a resource-constrained device.

Comprehensive analysis of four major rice diseases through the use of a freely available data set, which will show that automated annotation is a viable option to manual labeling in the detection of agricultural diseases.

2. Related Work

2.1. Deep Learning for Plant Disease Detection

The implementation of deep learning in the detection of plant diseases has developed quite quickly during the last ten years. Initial methods would mostly work with image classification, with models like VGG, ResNet, and Inception being used to classify all of a leaf image into disease classifications[6,7]. Although they are very accurate at classifying, these algorithms cannot determine the location of the disease lesions in the image, which restricts their diagnostic value. This weakness is overcome by object detection approaches that not only classify but also localize the disease areas. The Faster R-CNN was one of the initial architectures used in plant disease detection[8]. Next came single-stage detectors, especially the YOLO family, which became popular because of their better speed-accuracy tradeoff. Tang et al[9] suggested Pest-YOLO using YOLOv5 and efficient channel attention in real-time pest detection. The ECA attention mechanism was used by Ni et al[10] in rice pest and disease classification. Most recently, Jia et al[5] created MobileNet-CA-YOLO by substituting YOLOv7 backbone with MobileNetV3 and adding Coordinate Attention, reaching lightweight and precise detection.

2.2. Automated Annotation Methods

Minimizing the annotation workload has turned into a field of active research. Semi-supervised and weakly-supervised learning strategies utilize small quantities of labeled data along with large volumes of unlabeled data [11]. The Segment Anything Model (SAM) [12] allows zero-shot segmentation under prompt guidance, and Grounding DINO [13] provides open-vocabulary detection using text prompts. Nevertheless, these techniques tend to fail when it comes to domain-specific agricultural imagery where disease lesions have very slight visual properties that may be very different than ordinary object category. Segmentation methods based on colours are a less complex but effective solution to agricultural problems. The specific colors resulting due to plant diseases- yellowing, browning, or whitening of leaf tissues - can be quantified using the color space analysis in HSV, LAB, or YCbCr spaces[14]. Such methods are understandable, computationally inexpensive, and do not need training data so they can be used in automated annotation solutions.

2.3. YOLOv11 Architecture

The current version of the YOLO line by Ultralytics is YOLOv11 (2024), which has better feature extraction using C3k2 blocks, better feature fusion using SPPF and C2PSA modules and better detection heads. The nano edition (YOLOv11n) is very well-balanced in terms of accuracy and efficiency, with just 2.58M parameters and 6.3 GFLOPs, and thus it is especially applicable in edge deployment situations within precision agriculture.

3. Proposed Method

The suggested structure contains two major steps: (1) automated lesion annotation through color-space analysis and (2) detection of diseases based on YOLOv11n.

3.1. Automated Lesion Annotation

We use an automated annotation process which takes advantage of the fact that various rice diseases manifest as visibly different colour changes in the leaf tissue. Through examination of the HSV (Hue-Saturation-Value) colour space attributes of every disease class, we devise detection strategies specific to a given disease in order to isolate the areas with lesions and provide bounding box annotations.

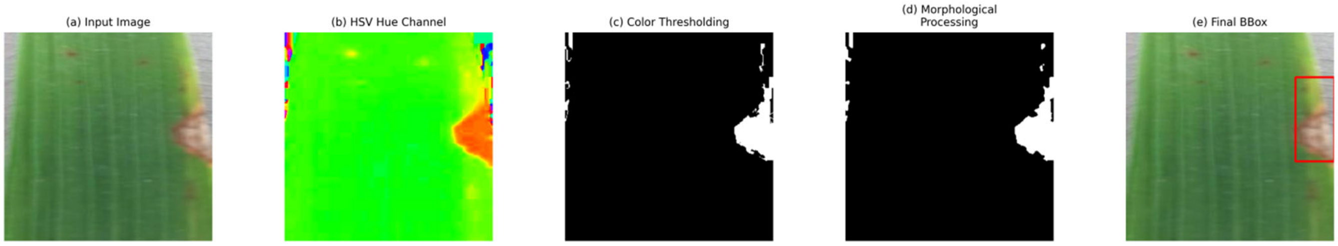
(a) Input Image
(b) HSV Hue Channel
(c) Color Thresholding
(d) Morphological Processing
(e) Final BBox

Figure 1. Annotation pipeline visualization

3.1.1. Preprocessing

For each input image, we perform the following preprocessing steps:

1. Color space conversion: Convert the BGR image to HSV and grayscale representations.
2. Foreground extraction: Intensity thresholding can be used to partition the leaf region into pure black ($V < 15$) and pure white ($V > 250$) backgrounds.

3.1.2. Disease-Specific Lesion Detection

We will discuss four custom-made detection methods depending on the pathological properties of every disease: Bacterial Blight (Xanthomonas oryzae): This disease causes pale gray-white streaks in leaf veins. To detect lesion regions we use:

- High Value channel threshold ($V > 160$)
- Low Saturation channel threshold ($S < 100$)
- Intersection with foreground mask to exclude background

Blast lesions caused by Rice Blast (Pyricularia oryzae) are spindle-shaped spots with a middle that is grayish-white and edges that are brown. Detection criteria:

- Low Saturation ($S < 120$) combined with moderate Value ($80 < V < 230$)
- Hue exclusion of healthy green range (H: 25–90°)
- Foreground constraint

Brown Spot (Bipolaris oryzae): Brown spot causes small, diffuse dark brown spots occupying very little area (<1% of the leaf). Our approach is based on local contrast:

The computation of the Gaussian blur subtraction is done by calculating the absolute difference between the original grayscale image and its Gaussian-blurred ($\sigma=10$, kernel=21x21) form.

- Dark spot detection: pixels where original intensity is >15 units below the blurred reference
- Hue anomaly detection: pixels deviating >15° from the green center ($H \approx 40°$)

Tungro (Rice Tungro Virus) Tungro leads to massive yellowing and orange coloring of leaves. Uses in detection:

- Yellow-orange Hue range (H: 8–35°) with moderate Saturation ($S > 40$)
- Exclusion of healthy green regions (H: 35–90°)
- Foreground constraint

3.1.3. Post-processing and Bounding Box Generation

After binary mask generation, we apply morphological operations:

1. Closing (elliptical kernel, 5×5): Connect adjacent lesion fragments
2. Opening (elliptical kernel, 3×3): Remove small noise artifacts
3. Contour extraction: Find connected components in the processed mask
4. Bounding box filtering:
   - Minimum area: 0.1–0.5% of image area (disease-dependent)
   - Maximum area: 85% of image area
   - Aspect ratio constraint: $\max(w,h)/\min(w,h) < 15$
5. Merging: If more than 15 boxes are detected, merge into

encompassing bounding rectangles

6. Backup: In the case of images that do not have enough detectable lesion pixels, a bounding box of 70% centered area is assigned

The output annotations are in YOLO form: <class_id> <x_center> <y_center> <width> <height> (normalized coordinates). Figure 2. Comparison between automatic annotation effects: original images (top row) and annotation results (bottom row) and the annotation performance on four types of rice leaf diseases.

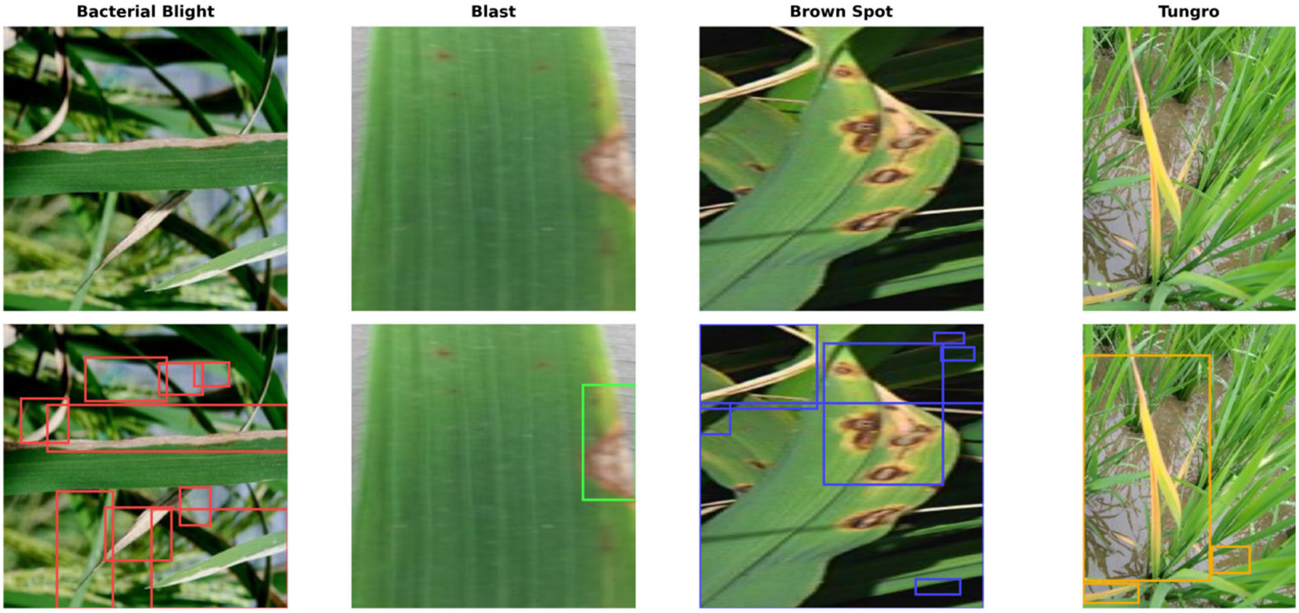

Figure 2. Comparison of automatic annotation effects

3.2. YOLOv11n Detection Model

YOLOv11n (nano) is chosen as the detection backbone because of its best trade-off between performance and computational efficiency. The architecture has:

- Backbone: CSPDarknet with C3k2 blocks for efficient multi-scale feature extraction

Neck: SPPF (Spatial Pyramid Pooling - Fast) and C2PSA (Cross-Stage Partial with Spatial Attention) as a means of improving feature fusion

Decoupled detector heads to classify and localize at three scales. Parameters: 2.58M (fused)

- FLOPs: 6.3G at 640×640 input resolution

The model is pre-trained using the COCO-weights and then fine-tuned on the automatically annotated rice disease dataset. The transfer learning technique enables the model to take advantage of general object detection expertise and customize it based on the domain-specific characteristics of diseases.

3.3. Training Configuration

Table 1 is a summary of the most significant training configuration parameters used in the experiment such as choosing the optimizer and its learning rate schedule, batch size, training epoch settings, and other fundamental hyperparameters that guarantee the reproducibility of the model training procedure.

Table 1. Key Training Configuration Parameters of Rice Disease Detection Model

Hyperparameter	Value
Input resolution	640 × 640
Batch size	16
Epochs	100
Optimizer	SGD (lr=0.01, momentum=0.937)
Learning rate schedule	Cosine annealing with warmup
Early stopping patience	20 epochs
Data augmentation	Mosaic, MixUp, HSV jitter, flip, scale
Random seed	42

4. Experiments

4.1. Dataset

To train our model, we will utilize the Rice Leaf Disease Image Samples dataset[15] which is freely available at Mendeley Data under the CC BY 4.0 license. It has a total of 5,932 images in 4 disease classes:

Table 2. Rice Leaf Disease Image Samples dataset overview

Category	Images	Description
Bacterial Blight	1,584	Gray-white streaks along veins
Blast	1,440	Spindle-shaped gray-white lesions
Brown Spot	1,600	Small scattered dark brown spots
Tungro	1,308	Extensive yellowing/orange discoloration
	Total	5932

The data set is divided randomly into training (80 percent, 4,745 images), validation (10 percent, 593 images), and test (10 percent, 594 images) sets using stratified sampling in order to preserve the proportions of classes (random seed = 42).

4.2. Automated Annotation Results

Table 3 presents the statistical results of automated annotation for four rice disease categories.

Table 3. Automated annotation results across disease categories.

Category	Total Images	Successful Detection	Fallback	Avg. Boxes/Image
Bacterial Blight	1,584	1,572 (99.2%)	12	3.2
Blast	1,440	1,398 (97. 1%)	42	2.8
Brown Spot	1,600	1,577 (98.6%)	23	6.8
Tungro	1,308	1,303 (99.6%)	5	4.6
Overall	5932	5,850 (98.6%)	82	4.4

This high successful detection rate (98.6 percent) confirms that our color space-based methodology is correct. The largest number of annotations per image (6.8) are created by Brown Spot because the lesions are scattered, whereas Blast has the least amount of annotations (2.8) because of the large and more unified lesion pattern.

4.3. Detection Performance

The trained YOLOv11n model has been tested on the held-out test set (594 images). The main conclusions are presented in Table 4.

Table 4. Per-class detection performance.

Disease	AP@0.5	AP@0.5:0.95	Instances (Test)
Bacterial Blight	0.9731	0.8449	551
Blast	0.9813	0.8669	616
Brown Spot	0.8556	0.6469	939
Tungro	0.9244	0.7427	629

Figure 3. Precision-Recall (PR) plot of the test set illustrating the precision/recall trade-off at various confidence limits.

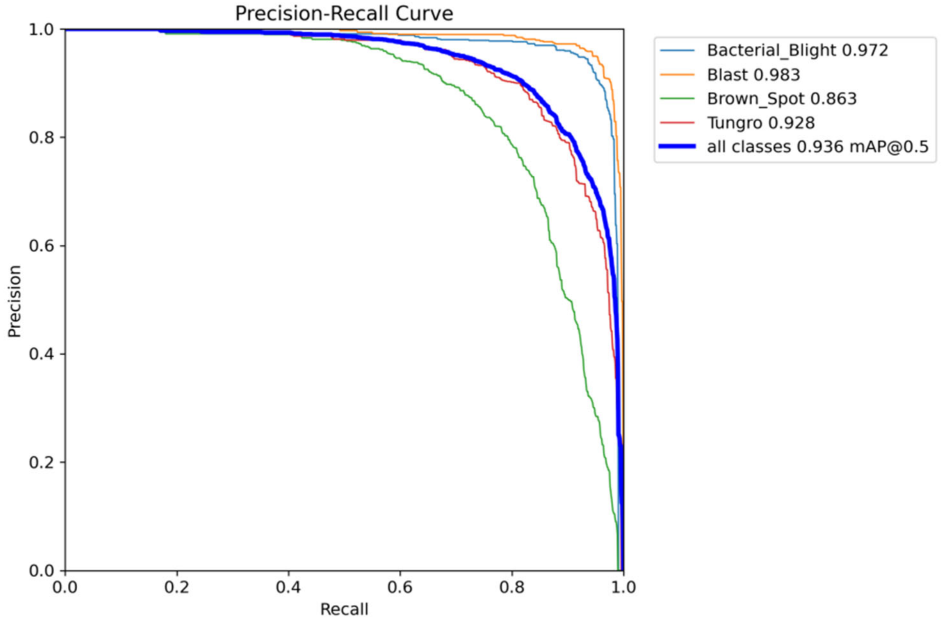

Figure 3. Precision-Recall (PR) curve of the test set

Figure 4 F1 -Confidence curve of the test set indicating change in F1 score with confidence threshold.

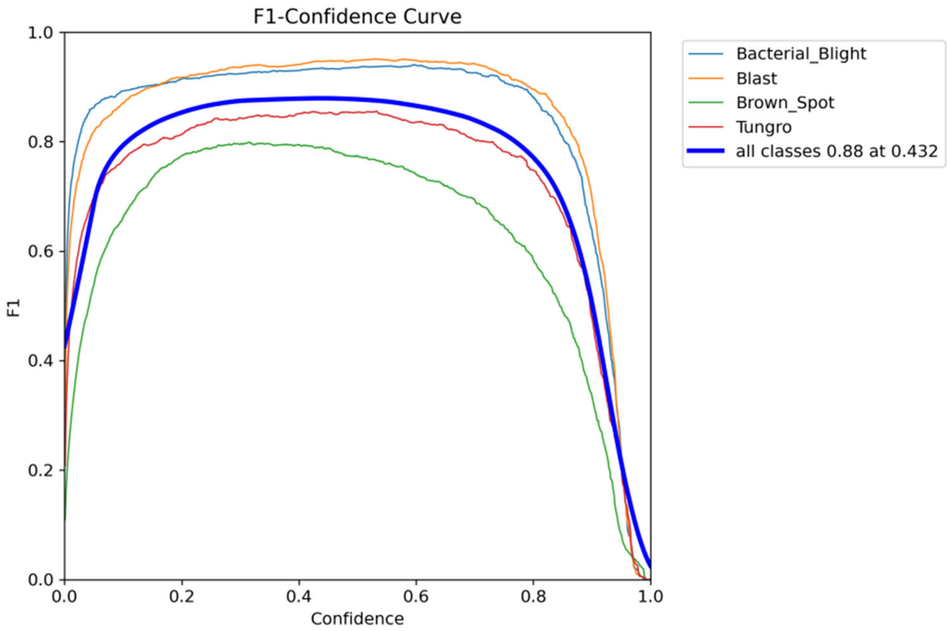

Figure 4. F1-Confidence curve of the test set

Figure 5. The ground truth annotation display of the test set showing the manual verification of bounding box labels of diseases lesions.

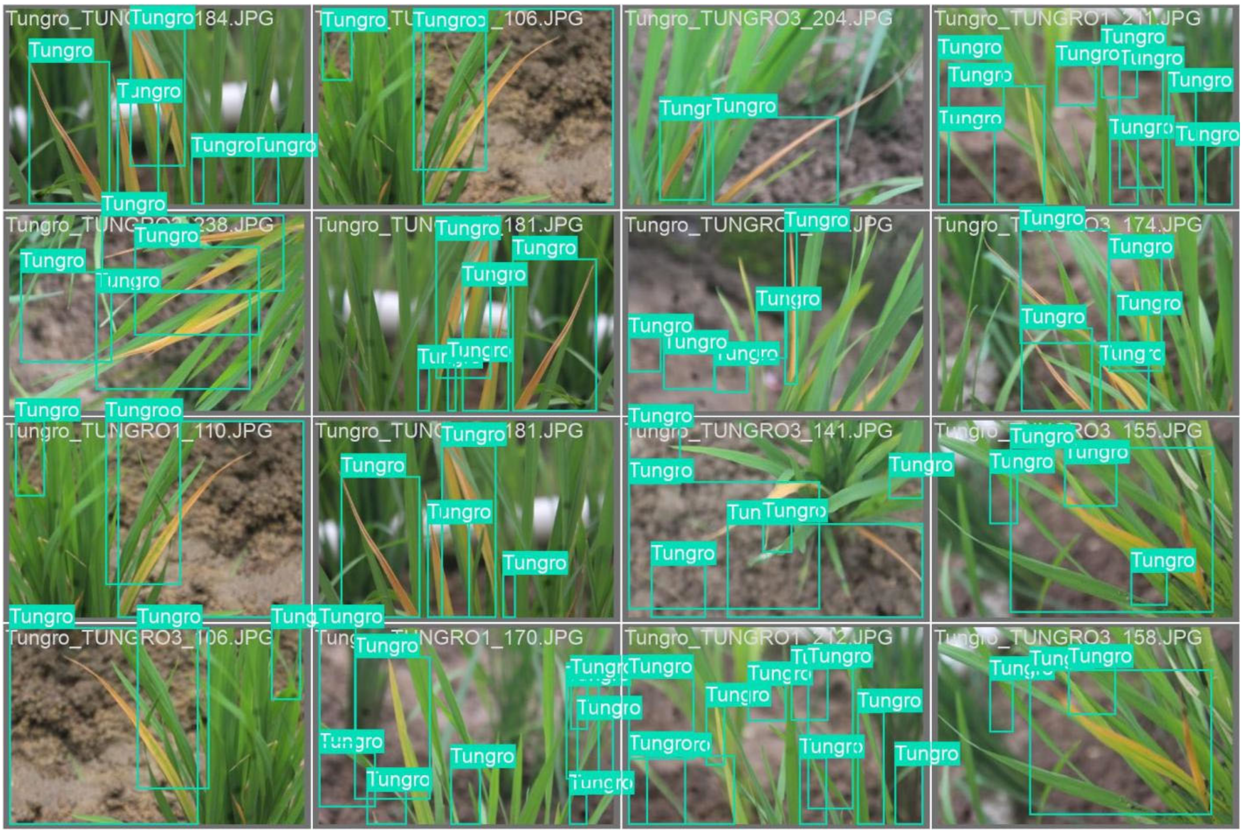

Figure 5. Ground truth annotation display of the test set

Image 6. The result of model prediction on the test set in comparison with the ground truth to illustrate detection quality.

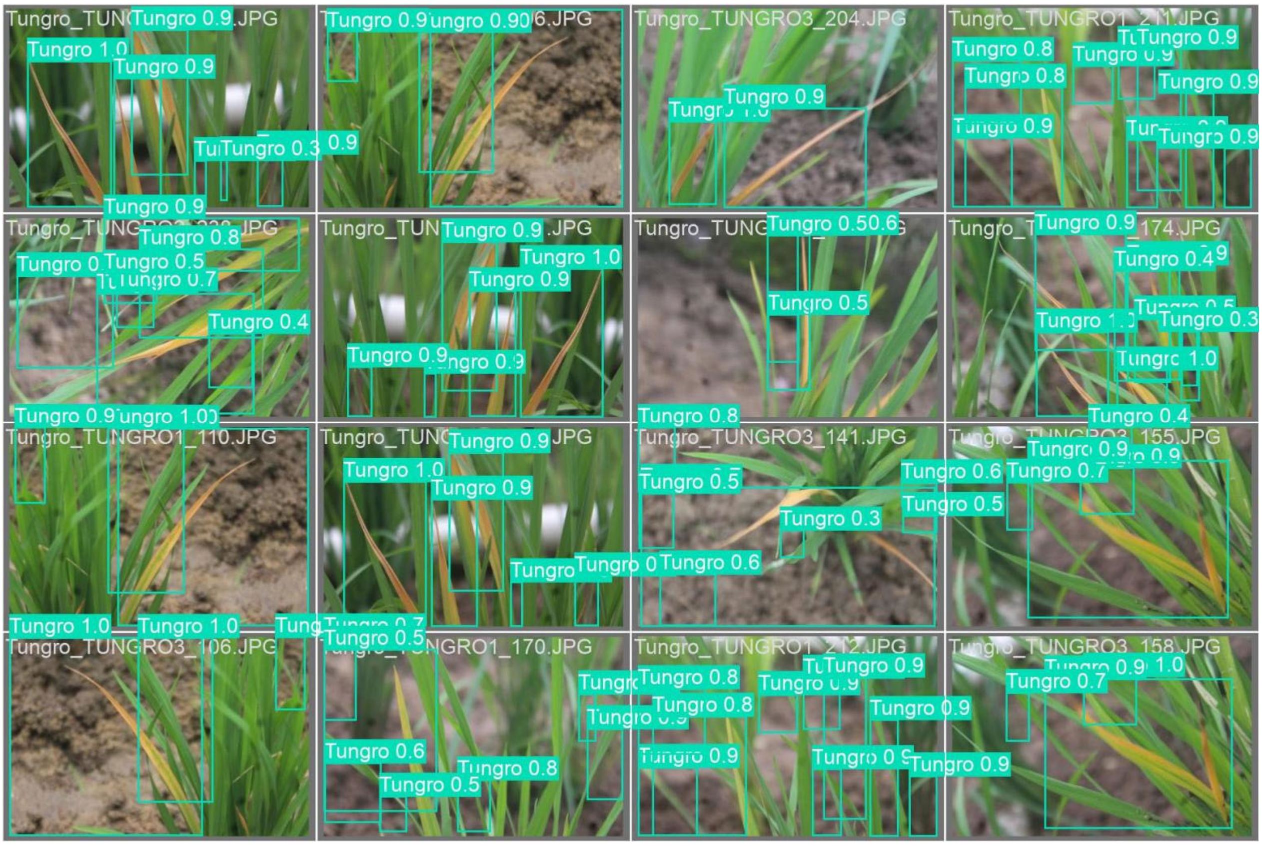

Figure 6. Model prediction result display of the test set

4.4. Model Efficiency

Table 5. Model efficiency comparison.

Model	Params (M)	FLOPs (G)	Size (MB)	Inference (ms)	mAP@0.5
MobileNet-CA-YOLO [5]	~8.5*	~16.2*	~17*	~12*	0.937*
YOLOv7 [16]	36.9	104.7	74.8	~7.0	—
YOLOv8n	3.2	8.7	6.3	~1.2	—
YOLOv11n (Ours)	2.58	6.3	5.2	1.0	0.934

*Values estimated from the original paper.

Comparison of the efficiency of the proposed YOLOv11n model with other state-of-the-art models (MobileNet-CA-YOLO, YOLOv7, YOLOv8n). Our YOLOv11n model has similar performance in mAP@0.5 (0.934 compared to 0.937) as MobileNet-CA-YOLO but is much more light-weighted: 70 times fewer parameters, 61 times fewer FLOPs, and 69 times smaller model size. A speed of 1.0ms per image when using an RTX 4090 GPU shows the high potential of real time applications.

4.5. Training Dynamics

The model was trained to the full 100 epochs without early stopping, which implies that learning was steady across the training process. Figure 7. Training curves: changes in loss, mAP, Precision and Recall with training epochs, demonstrating the convergence and performance improvement of the model in training. Main findings:

- Convergence: Training losses (box, classification, and DFL) showed steady decrease throughout all 100 epochs
- Best validation performance: Achieved around epoch 80 with mAP@0.5 = 0.936
- No overfitting: Validation metrics remained stable in the final epochs, indicating good generalization
- Training time: Approximately 27 minutes on a single NVIDIA RTX 4090 GPU

4.6. Per-Class Analysis

The performance of detection depends on the type of disease because it is determined by the visual properties of every disease:

Blast (AP@0.5 = 0.981): The best detection rate was reached. Blast injuries are characterized by typical spindle shapes, which are pale gray-white in color in the center and dark brown on the edges, giving a very good visual clue of both the automated annotations and model detection.

Bacterial Blight (AP@0.5 = 0.973): The gray-white water-soaked streaks on leaf veins form consistent patterns which are very well captured by the high-brightness, low-saturation detection strategy.

Tungro (AP@0.5 = 0.924): Large areas of yellowing are caused by tungro virus creating large areas that are easy to find. Nevertheless, the vague interface between healthy and diseased tissue somewhat decreases the accuracy of localization.

Brown Spot (AP@0.5 = 0.856): The worst performance out of the four classes due to: (1) very low lesion area size (less than 1 percent of the image area); (2) high density of scattered spots leading to difficulty in annotating them; and (3) delicate color difference between brown spots and the black green leaves.

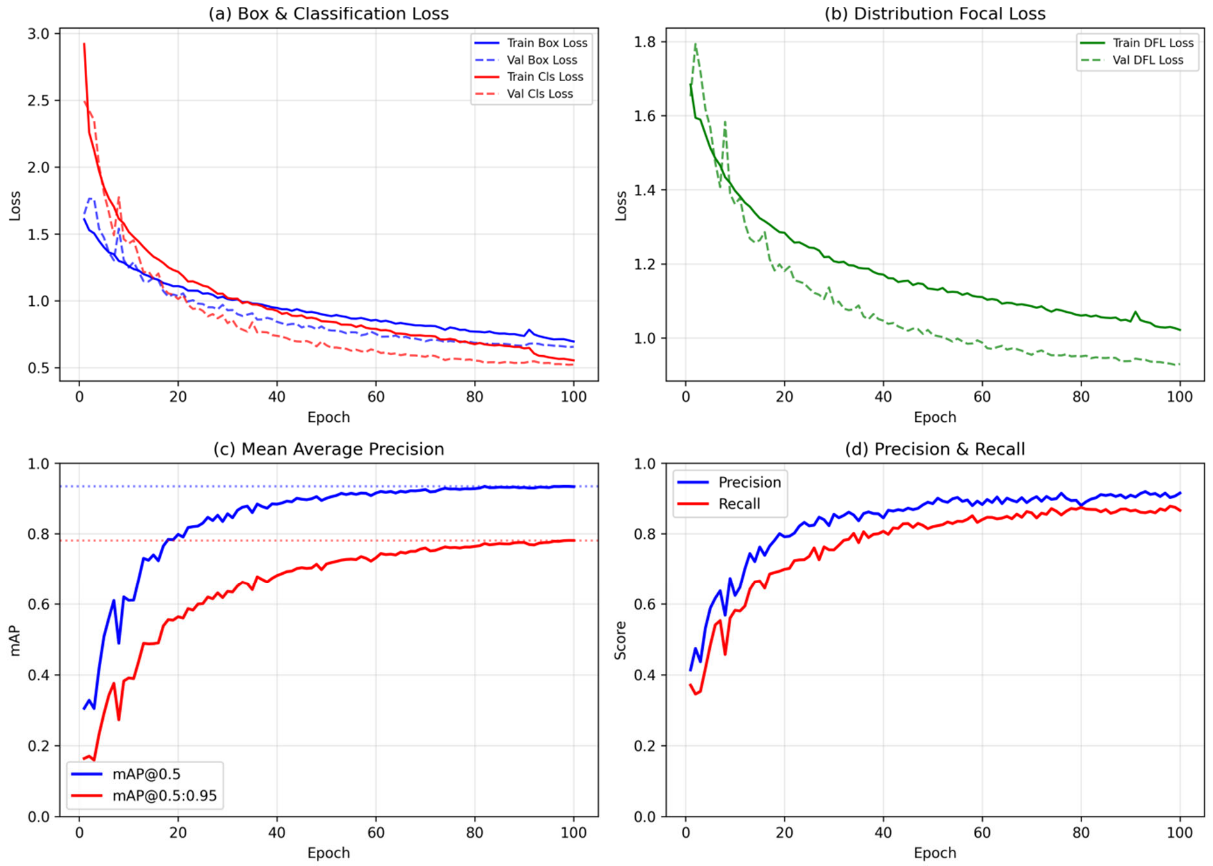

Figure 7. Training curves

5. Discussion

5.1. Advantages of the Automated Annotation Approach

The suggested annotation based on colors has a number of benefits in comparison with manual annotation and general-purpose foundation models (e.g., SAM, Grounding DINO):

Domain specificity: Using pathological information on the colour signs of a disease, our strategy is able to cover a 98.6 percentage without any training information or optimal model tuning.

Interpretability: All the annotation decisions may be mapped to particular color space thresholds, which allows agronomists to check and revise the criteria of annotations.

Computational efficiency: Annotation is performed by using simple image processing procedures (color conversion, thresholding, morphological procedures) and processes about 100 images per second on an average CPU.

4. Zero annotation cost: There is no need to do any human annotation work; hence it removes the major bottleneck of creating detection datasets in agriculture.

5.2. Limitations and Future Work

Several limitations merit discussion:

Annotation quality ceiling: Although the automatically created bounding boxes have reached 98.6 percent coverage, they might not be able to accurately demarcate borders of a lesion especially when it is irregularly shaped or overlaps other lesions. There is a noise ceiling on the attainable detection rate.

2. Brown Spot challenge: The very tiny nature of brown spot lesions makes it difficult to detect them on both annotation (which needs aggressive local contrast methods) and detection (small object detection is still an open problem in computer vision).

3. Color-dependency: The method is based on the assumption that a color of disease lesions is constant, but it can be affected by light conditions, camera settings or the stage of disease development. Adaptive thresholding or lighting normalization are areas of future work to explore.

Single-disease assumption: It is presumed that each image has only one kind of disease as per the structure of the dataset. The extension to multi-disease co-occurrence situations would necessitate more logic.

Promising directions for future improvement include:

Refinement of iteration: Starting with initial auto-annotations to train a detection model and then using the model predictions to improve annotations (pseudo-labeling loop)

Multi-scale detection: Use of bigger input dimensions (1280x1280) or specific small object detection architectures on Brown Spot

- Cross-dataset validation: Evaluating generalization to field-captured images under diverse environmental conditions

5.3. Practical Implications

The proposed framework has significant practical implications for precision agriculture:

Fast dataset creation: Agronomists are able to rapidly create detection datasets on new crops diseases by formulating color detection rules specific to the diseases, avoiding the months of manual annotation.

Edges deployment: The 5.21MB model size and 1.0ms inference rate allows running on mobiles, drones, and IoT devices to monitor in field real time.

Scalability: The automated pipeline can easily be scaled to other crops and diseases by changing the color space detection parameters to the particular pathological features.

6. Conclusion

The current paper is an automated annotation and detection system of rice leaf disease which is aimed at solving one of the most critical issues of the annotation cost in the context of agricultural object detection. Our procedure uses the color space features of a disease in the HSV domain to automatically create bounding box annotations with 98.6 percent coverage of 5,932 images that cover four leading rice diseases. The YOLOv11n model with only 2.58M parameters and 1.0ms inference time is trained on these auto-annotations, achieving 93.36% mAP@0.5 and 77.53% mAP@0.5:0.95 on the test set.

Our findings indicate that: (1) automated color space-based annotation is a feasible and cost-effective option of manual labeling in detecting agricultural diseases; (2) the lightweight YOLOv11n can be as effective as more sophisticated architectures at substantially lower cost; and (3) various rice diseases have different levels of detection difficulty with Brown Spot being the most difficult to detect because it has very tiny lesions

The framework proposed provides a realistic way to develop crop disease detection systems with little human intervention, which will result in the development of precision agriculture and automated monitoring of crop health.

References

[1] Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., & Nelson, A. (2019). Global impact of pathogens and pests on important food crops. *Nature Ecology & Evolution*, *3*(3), 430–439. https://doi.org/10.1038/s41559-019-0833-6.

[2] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 779–788).

[3] Liu, J., & Wang, X. (2021). Plant diseases and pests detection on basis of deep learning: Review. *Plant Methods*, *17*, 22. https://doi.org/10.1186/s13007-021-00722-9

[4] Li, L., Zhang, S., & Wang, B. (2021). Plant disease detection and classification with the help of deep learning: A review. *IEEE Access*, *9*, 56683–56698. https://doi.org/10.1109/ACCESS.2021.3072807

[5] Jia, L., Wang, T., Chen, Y., Zang, Y., Li, X., Shi, H., & Gao, L. (2023). MobileNet-CA-YOLO: A new version of the YOLOv7 model that is based on the combination of the MobileNetV3 and attention mechanism to detect rice pests and diseases. *Agriculture*, *13*(7), 1285. https://doi.org/10.3390/agriculture13071285

[6] Mohanty, S. P., Hughes, D. P., & Salathé, M. (2016). Deep learning applied to image based detection of plant diseases. *Frontiers in Plant Science*, *7*, 1419. https://doi.org/10.3389/fpls.2016.01419

[7] Ferentinos, K. P. (2018). The models of deep learning in detecting and diagnosing plants with diseases. *Computers and Electronics in Agriculture*, *145*, 311–318. https://doi.org/10.1016/j.compag.2018.02.016

[8] Fuentes, A., Yoon, S., Kim, S. C., & Park, D. S. (2017). Robust deep-learning-based detector of plant diseases and pests recognition in real time based on tomato plants. *Sensors*, *17*(9), 2022. https://doi.org/10.3390/s17092022

[9] Tang, Z., Lu, J., Chen, Z., Qi, F., & Zhang, L. (2023). Enhanced Pest-YOLO: A real-time pest detector that works on the basis of the effective channel attention mechanism and transformer encoder. *Ecological Informatics*, *78*, 102340. https://doi.org/10.1016/j.ecoinf.2023.102340

[10] Ni, H., Shi, Z., Karungaru, S., Lv, S., Li, X., Wang, X., & Zhang, J. (2024). The classification of major pests and diseases of rice based on the ECA attention mechanism. *Agriculture*, *14*(5), 682. https://doi.org/10.3390/agriculture14050682

[11] Sohn, K., Berthelot, D., Carlini, N., Zhang, Z., Zhang, H., Raffel, C. A., & Li, C. L. (2020). FixMatch: Semi-supervised learning made simple through consistency and confidence. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 596–608).

[12] Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., & Gustafson, L. (2023). Segment anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4015–4026).

[13] Liu, S., Zeng, Z., Ren, T., Li, F., Zhang, H., & Yang, J. (2024). Grounding DINO: Marrying DINO with grounded pre-training to detect open-set objects. In *Proceedings of the European Conference on Computer Vision* (pp. 38–55).

[14] Barbedo, J. G. A. (2019). The identification of plant diseases based on individual lesions and spots with the help of deep learning. *Biosystems Engineering*, *180*, 96–107. https://doi.org/10.1016/j.biosystemseng.2019.02.004

[15] Sethy, P. K., Barpanda, N. K., Rath, A. K., & Behera, S. K. (2020). Support vector machine based identification of diseases in rice leaves by deep features. *Computers and Electronics in Agriculture*, *175*, 105527. https://doi.org/10.1016/j.compag.2020.105527

[16] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies becomes a new state-of-the-art in real-time object detector. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7464–7475). https://doi.org/10.1109/CVPR52729.2023.00719