Grass-weed Identification Using Neural Networks in the Rice Cultivar

Yolov8 based model to classify rice and weed species in high resolution images

Student: -

Student ID: -

University

Table of contents

[1. Introduction 3](#_Toc196065519)

[2. Literature Review 4](#_Toc196065520)

[3. Methodology 7](#_Toc196065521)

[Data Collection 7](#_Toc196065522)

[Data Preprocessing 7](#_Toc196065523)

[Model Selection 10](#_Toc196065524)

[Model Architecture 10](#_Toc196065525)

[Training and validating 13](#_Toc196065526)

[Parameter Optimization 14](#_Toc196065527)

[Implementation tools 15](#_Toc196065528)

[Implementation software tools 16](#_Toc196065529)

[Experimental Design and Model Validation 17](#_Toc196065530)

[4. Results 18](#_Toc196065531)

[6. Discussion and conclusion 22](#_Toc196065532)

[Classification Performance & Challenges 22](#_Toc196065533)

[Annotation & Prediction Discrepancies 22](#_Toc196065534)

[Key Observations 23](#_Toc196065535)

[Practical Applications for Farmers 23](#_Toc196065536)

[7. References 25](#_Toc196065537)

[Appendices 26](#_Toc196065538)

[Appendix A: Images included in training detection model 26](#_Toc196065539)

[Appendix B: Model settings used for image size 640 28](#_Toc196065540)

[Appendix C: Model settings used for image size 1200 29](#_Toc196065541)

[Appendix D: Classification performance 30](#_Toc196065542)

[Appendix E: Annotations and predictions 32](#_Toc196065543)

### 1. Introduction

Rice cultivation plays a vital role in both global and European agriculture, significantly contributing to food security, rural economies, and economic sustainability. In Europe, rice production spans approximately 500,000 hectares, primarily across Italy, Spain, Greece, Portugal, and France, which require precise irrigation and weed management to achieve high yields of up to 10 tonnes per hectare. Despite being a relatively small portion of global rice production, European rice farming faces key challenges such as herbicide-resistant weeds, rising production costs, and environmental sustainability concerns.

Weeds are a major obstacle in rice fields, competing with crops for water, light, and nutrients, ultimately leading to reduced yields and higher operational costs. Traditional weed management approaches, which rely heavily on herbicide application, have become less effective due to increasing herbicide resistance. This increasing resistance necessitates innovative, technology-driven solutions for effective weed detection and management.

This study introduces a neural network-based approach using the YOLOv8 model, designed to enhance weed identification in rice fields through high-resolution imagery and machine learning techniques. Experimental results demonstrated that image resolution and validation settings significantly influence classification accuracy. Models trained with larger images (1200 pixels) consistently outperformed those trained on smaller image sizes (640 pixels), achieving higher precision and recall scores. Specifically, models using image size 1200 achieved an mAP50 of ~0.4 and an mAP50-95 of ~0.25, compared to mAP50 values of 0.18-0.22 and mAP50-95 values of 0.07-0.1 in models trained on smaller images.

Furthermore, classification errors were often observed when large overlapping weed groups were present, making accurate identification more difficult. 1ECHG – Weed emerged as the most frequent misclassification for ORYSA and ORYSN, highlighting the need for refined annotation techniques, better resolution imagery, and improved feature extraction methods. Additionally, background predictions in YOLO models affected classification accuracy by mis categorizing certain objects as background. Post-processing techniques aimed at reducing background-labeled predictions helped improve overall model performance.

Artificial intelligence (AI) driven agricultural solutions enable autonomous crop monitoring, yield forecasting, and real-time weed detection, facilitating smart farming practices. Farmers can leverage IoT sensors, drone imaging, and robotic automation to enhance precision agriculture, reducing labor costs while improving sustainability.

This research underscores the importance of optimized image-based weed detection, examining annotation methodologies, classification accuracy, and data augmentation strategies. With an emphasis on real-world agricultural applications, this study offers insights into how AI-powered weed detection can improve economic efficiency, promote sustainable rice cultivation, and ensure long-term viability in European rice farming.

### 2. Literature Review

Rice cultivation plays a pivotal role both globally and within Europe, significantly contributing to food security and rural livelihoods. Globally, rice is a staple food for more than half the population, providing vital nutritional support. Approximately 158 million hectares are utilized worldwide for rice farming, yielding over 700 million tonnes annually. However, traditional practices such as flooded paddies result in methane emissions, which account for 12% of global agricultural greenhouse gas outputs, posing a major challenge in mitigating climate change (Chauhan, 2012; Ziska et al., 2015).

In Europe, rice is grown across approximately 500,000 hectares, predominantly in countries like Italy, Spain, Greece, Portugal, and France. The cultivation primarily focuses on *Japonica* rice varieties, which are irrigated to achieve yields as high as 10 tonnes per hectare. Despite contributing only 0.4% of global rice production, European rice farming plays an essential role in agricultural gross domestic product (GDP), rural economies, and food security. However, challenges such as weed management, water scarcity, high production costs, and the necessity for sustainable farming practices demand innovative solutions (European Commission, 2025; Sustainable EU Rice Project, 2025).

Weeds significantly hinder rice production by competing with crops for essential resources such as light, water, and nutrients, thereby reducing yields and quality. They also increase production costs through labor and chemical interventions required for their management. Research on eco-economic thresholds has highlighted the necessity of integrated weed management strategies, which balance environmental sustainability with economic viability (Tian et al., 2020). Additionally, herbicide resistance among weeds, especially in \_*Echinochloa spp*., has emerged as a critical challenge, emphasizing the need for innovative control methods (Godar and Norsworthy, 2023; Vulchi et al., 2024).

Certain weed species, such as *Echinochloa crusgalli*, *Cyperus difformis*, and Weedy (red) rice, are particularly problematic for rice production:

*Echinochloa crusgalli*: This highly competitive weed reduces yields and complicates herbicide efficacy due to its adaptability (Godar and Norsworthy, 2023).

*Cyperus difformis*: Its prevalence has notable effects on rice yields and eco-economic thresholds, highlighting the importance of targeted management practices (Tian et al., 2020).

Weedy (red) rice: Known for its genetic complexity and adaptability, this weed lowers rice quality and poses significant challenges for sustainable farming worldwide (Ziska et al., 2015; Nadir et al., 2017).

The incorporation of machine learning (ML), AI, and computer vision into agriculture provides innovative opportunities to address challenges in rice cultivation. These technologies are particularly useful for early weed detection and yield prediction. For example, multi-spectral UAV data has proven effective in early weed mapping, enabling targeted interventions during initial growth stages (Stroppiana et al., 2018). Similarly, advanced computer vision methods are enhancing precision agriculture by distinguishing weeds from crops based on spectral and spatial patterns (Wu et al., 2021).

While promising, these technologies face several challenges. High-quality datasets and computational capabilities are critical to ensure accuracy and scalability. Furthermore, environmental variability, such as differences in climatic or soil conditions, complicates the reliability of predictive models for managing weeds like *Echinochloa spp*., which exhibit diverse germination patterns influenced by ecological factors (Yuan et al., 2025). Despite these limitations, ML and AI offer substantial potential to boost agricultural sustainability and productivity (Botero-Valencia et al., 2025; Araújo et al., 2023).

While considerable progress has been made in understanding the global and regional impacts of rice cultivation, this thesis focusses on providing an effective computer vision-based detection model, emphasizing practical and cost-effective adaptation for real-world agricultural implementation and sustainability.

Computer Vision in Weed Management: Further research is needed to develop advanced image-based solutions for sustainable weed control, targeting weed species that cause significant ecological and economic damage in agriculture.

Data Quality and Data Usability: Enhanced data collection frameworks are crucial for optimizing AI and ML applications in weed detection, yield forecasting, and environmental sustainability, ensuring accuracy, accessibility, and cost-effective practical implementation in agricultural decision-making.

Socio-economic Studies: Research on the economic impact of adopting innovative technologies in smallholder farming systems should also address usability and affordability, ensuring practical implementation and cost-effectiveness for farmers while bridging the gap between innovation and accessibility.

### 3. Methodology

## Data Collection

Annotated high-resolution images of plants and weeds were provided by Elgo Dimitra under the guidance of Dr. G. Peteinatos, camera specifications and used settings shown in Table 1.

|  |  |
| --- | --- |
| Specifications | |
| Make | Xiaomi |
| Model | 23090RA98G |
| Software | MediaTek Camera Application |
| lens size | 23 mm |
| optical format | 1/1.4 |
| pixel size | 0.56 µm |
| shutter | electronic rolling shutter |
| focus | multi-directional Phase Detection Auto Focus (PDAF) |
| Settings | |
| shutter speed | 1/100s |
| ISO | automatically calibrated |

Table : Camera specifications and settings

The dataset included species identified using EPPO codes: Orysa (*Oryza sativa*), Orysn (*Oryza sativa Indica*), Cypdi (*Cyperus difformis*), and 1echg (*Echinochloa*).

## Data Preprocessing

* + Image Selection: Images with sufficient resolution were selected for further processing.

The selection of the images for model training were based on the resolution of the provided images, from the initial dataset which was provided the excluded images and rationale are shown in Table 2. Detailed data about pictures included in the model's training can be found in Appendix A.

|  |  |  |  |
| --- | --- | --- | --- |
| Filename | Width | Height | reason for exclusion |
| 1729767882460.jpg | 1200 | 1600 | resolution to low |
| 1729767882491.jpg | 1200 | 1600 | resolution to low |
| 1729767882506.jpg | 1200 | 1600 | resolution to low |
| 1729767882529.jpg | 1200 | 1600 | resolution to low |
| 1729767882548.jpg | 1200 | 1600 | resolution to low |
| 1729767882567.jpg | 1200 | 1600 | resolution to low |
| 1729767882595.jpg | 1200 | 1600 | resolution to low |
| IMG\_20241029\_103212 (1).jpg | 2296 | 4080 | duplication |
| IMG\_20241029\_124651 (1).jpg | 2296 | 4080 | duplication |
| IMG\_20241029\_125158 (1).jpg | 4080 | 2296 | duplication |
| IMG\_20241029\_125206 (1).jpg | 4080 | 2296 | duplication |
| IMG\_20241029\_125217 (1).jpg | 4080 | 2296 | duplication |
| IMG\_20241029\_125222 (1).jpg | 4080 | 2296 | duplication |
| IMG\_20241029\_125226 (1).jpg | 4080 | 2296 | duplication |

Table : Excluded images with rationale

* + Image segmentation

To ensure a robust and unbiased evaluation, all images were systematically partitioned following the structure shown in figure 1. Due to the limited number of original images, each image corresponds to a distinct growth stage within the plant or weed life cycle, allowing for comprehensive coverage of developmental variations. Additionally, careful dataset curation was performed to guarantee that no training images are present in the validation set, ensuring reliable performance assessment and minimizing data leakage risks. This approach enhances the integrity of classification and detection tasks while maintaining consistency in image segmentation for accurate model development.

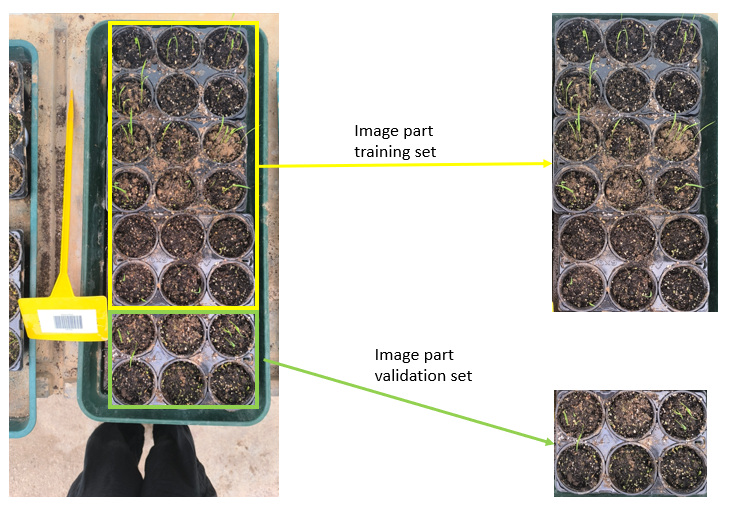
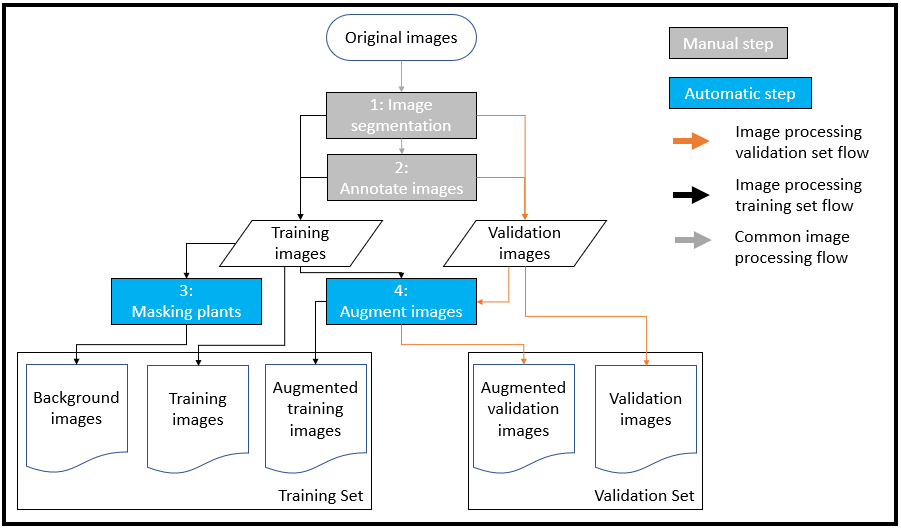


Figure : Split images (IMG\_20241021\_105827)

* + Annotation: Selected images were annotated to label plants and weeds for training and validation purposes, see figure 2, 2: Annotate images.

LabelImg (v1.8.1), an open-source tool for annotating images, was used to create bounding boxes required for object detection workflows (Human Signal, n.d.). These annotations were saved in the YOLO format to ensure compatibility with model training and assessment processes.

 Figure : Flowsheet image processing for training and validation set

* + Dataset Augmentation:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | augmentation percentage | | | | |
| threshold | 1ECHG - Weed | ORYSA | CYPDI - Weed | ORYSN - Weed | total training set |
| Original number annotations | 775 | 473 | 302 | 199 |  |
| 500 | 39.22% | 51.39% | 62.34% | 71.53% | 53.35% |
| 1500 | 65.93% | 76.03% | 83.24% | 88.29% | 77.43% |
| 2500 | 76.34% | 84.09% | 89.22% | 92.63% | 85.11% |

Table : Augmentation percentage for training datasets

The dataset was expanded through artificial augmentation to ensure that each classification group met a predefined instance threshold, resulting in an increased dataset containing a high degree of augmented images as can be seen in table 3.

Augmentation was also used to balance the dataset, reducing the potential for bias across diverse groups (see Figures 2 and 4: Augmented Images). Two augmentation techniques were applied to increase the number of images available for model training.

Random rotation, varying from -80° to 80°, was used to diversify image orientations.

Background masking, applied with a 50% probability, replacing background information with a random greyscale value. The masking area was determined by filtering pixels outside the HSV range. Specifically, for images containing *Cyperus difformis*, the boundary values were set at H = 25-75, S = 135-255, V = 50-255, whereas for *Oryza sativa*, *Oryza sativa Indica*, and *Echinochloa*, the range was H = 25-75, S = 50-165, V = 50-255. The background mask was generated by replacing pixels within the boundary area with black pixels.

Images were augmented until a threshold specified

## Model Selection

The YOLOv8 model a Convolutional Neural Network (CNN), was chosen for its speed and accuracy in object detection tasks. In agricultural applications, YOLOv8 can be used in identifying crops, pests, and diseases with high precision, enabling activities such as precision weed management, pest detection, and crop health monitoring. Its simplicity allows trained models to be deployed on devices with limited computational resources, making it suitable for field use. Furthermore, its real-time capabilities empower farmers to make data-driven decisions effectively.

## Model Architecture

YOLO Training Process

The training phase involves configuring YOLO with augmentation techniques, batch size adjustments, and optimization strategies to improve generalization. Specific enhancements, such as random rotations, background modifications, and bounding box refinements, are integrated to increase robustness. The dataset undergoes pre-processing, including image resizing, annotation corrections, and balanced augmentation, ensuring comprehensive coverage of diverse object instances. The training is performed over multiple epochs, leveraging Adam optimization, pseudocode for training yolo model is shown in figure 3.

Once training is completed, yolo validation is conducted using the validation dataset and the settings used for training the model, pseudocode for training shown in figure.

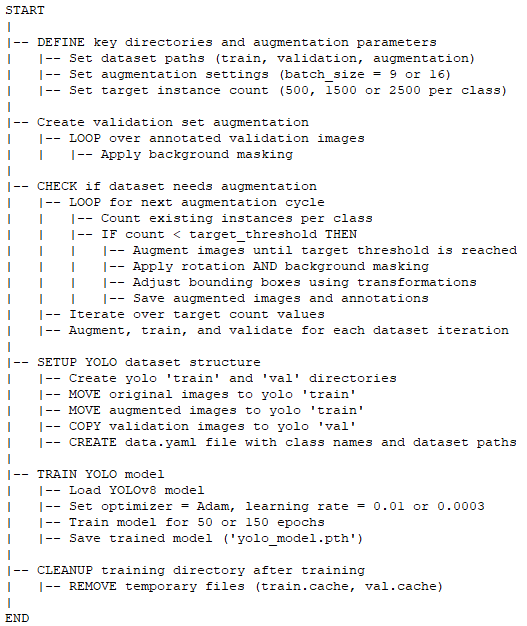


Figure : Pseudocode training detection model

Improve Classification Accuracy with Yolo Validation

Yolo validation is executed with alternative confidence threshold and Intersection over Union (IoU) setting using the validation dataset, pseudocode for classification improvement program is shown in figure 4. The aim is to improve classification results. This step involves computing key performance metrics, such as:

mAP50 (Mean Average Precision at IoU threshold 0.50)

mAP50-95 (Mean Average Precision across multiple IoU thresholds)

Recall (Ability to detect true objects)

Precision (Ratio of correctly classified objects among detected instances)

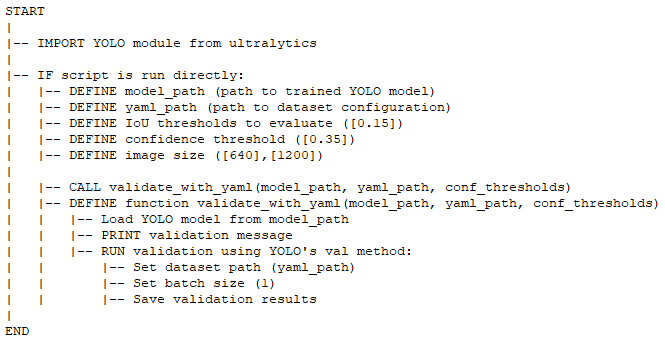


Figure : Pseudocode improving classification

Statistical Comparison Using Mann-Whitney U Test

To verify whether training and post training modifications lead to statistically significant improvements, Pairwise Mann-Whitney U Tests are performed. Given the non-parametric nature of this test, it is particularly useful when the dataset size is small, or normality assumptions are uncertain. The statistical analysis compares classification results across different training configurations, assessing whether performance enhancements yield meaningful improvements in object recognition, pseudocode for visualization and statistical calculations is shown in figure 5.

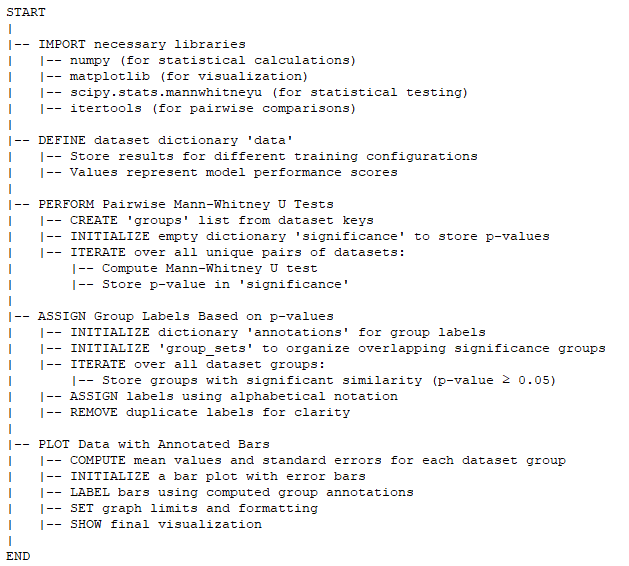


Figure :Pseudocode statistical comparison and visualization

Results are grouped based on mAP, recall, and precision metrics, ensuring that alterations in training strategies demonstrate statistically valid outcomes. Additionally, visual plots are generated to illustrate the impact of augmentation techniques on classification accuracy.

## Training and validating

To ensure a well-balanced training set, the dataset was curated to maintain an equal distribution of instances across all classification groups. This approach minimizes bias and improves model generalization. The training process was conducted with two different image sizes:

Image size = 640: Training was performed over 150 epochs.

Image size = 1200: Training was conducted for 50 epochs.

These settings were determined empirically, with the primary objective of ensuring stability in model convergence. The selection of epoch counts was based on monitoring training loss metrics—specifically, train/box\_loss, train/cls\_loss, and train/dfl\_loss—and comparing them against the corresponding validation metrics (val/box\_loss, val/cls\_loss, val/dfl\_loss). The finalized training strategy aimed to prevent excessive divergence between training and validation losses, ensuring a well-regularized model capable of performing consistently across unseen data.

## Parameter Optimization

## The model's hyperparameters tuning to improve the performance:

* + Optimizer: Adam optimizer was selected for its adaptability and stability:

- Faster convergence during training.

- Adjustment of learning rates per parameter, enhancing performance on diverse features.

- Stabilized weight updates, reducing oscillations.

* + Image size (imgsz): Initial size 640 and increased to 1200 to enable the model to capture finer details.
  + Batch size (batch): Reduced to 9 due to computational constraints caused by the increased image size.
  + Dropout: Set to 0.35 to minimize overfitting on the dataset.
  + Loss balancing (`dfl`): Decreased to 3 to balance box and classification losses, reducing risk of premature convergence.
  + Initial learning rate (`lr0`): Set to 0.0003 for gradual, precise weight updates, improving generalization on complex datasets and reducing overfitting.
  + IoU threshold: Set to 0.85 to achieve:

- Tighter bounding boxes, enhancing detection accuracy.

- Fewer false positives, improving localization quality.

- High-quality outputs critical for tasks like weed and plant detection in precision agriculture.

* + Confidence threshold (conf): Decreased to 0.25 to capture hard-to-detect objects, such as weeds and plants.
  + Patience: a hyperparameter used in early stopping, which helps improve training efficiency and prevent overfitting increased to 100

## Implementation tools

* + ASUS TUF Gaming A17 FA706IU\_FX706IU

- OS Name: Microsoft Windows 11 Home

- Version: 10.0.26100 Build 26100

- System Manufacturer: ASUSTeK COMPUTER INC.

- Processor: AMD Ryzen 7 4800H with Radeon Graphics, 2900 MHz, 8 Core(s), 16 Logical Processor(s)

- BIOS Version/Date: American Megatrends Inc. FA706IU.316, 12/03/2021

- Installed Physical Memory (RAM): 16.0 GB

- Total Physical Memory: 15.4 GB

- Total Virtual Memory: 49.4 GB

* + Xiaomi Redmi Note 13 Pro

- Processor: Qualcomm Snapdragon 7s Gen 2 (4nm)

- Main Camera: 200MP (wide) + 8MP (ultrawide) + 2MP (macro), 4K video recording

- Front Camera: 16MP, 1080p video recording

- Operating System: Android 13 with HyperOS

Image Acquisition Setup:

For image collection, a Xiaomi Redmi Note 13 Pro equipped with a 200-megapixel RGB camera (ISOCELL HP3, Samsung Semiconductors, Suwon-si, South Korea) was used. The standard MediaTec Camera Application was employed with default settings to ensure consistency across all captured images.

Camera Specifications & Settings:

The camera features a 23 mm lens, a 1/1.4" optical format, 0.56µm pixel size, an electronic rolling shutter, and multi-directional Phase Detection Auto Focus, ensuring high-precision image acquisition. Images were stored in JPEG format with a resolution of 4080 x 2296 pixels. A shutter speed of 1/100s was applied, while the ISO was automatically adjusted to optimize image quality under varying lighting conditions during measurements.

## Implementation software tools

* + OpenCV (cv2): Used for advanced image processing tasks, including extracting features and performing complex transformations on visual data.
  + Pandas: Efficiently organize, analyze, and manipulate tabular datasets, ensuring seamless data handling throughout the project.
  + NumPy: For numerical operations, enabling computations on large arrays and matrices.
  + Matplotlib and matplotlib.patches: Enabling the creation of detailed visualizations, such as bounding box overlays, to aid in interpreting results effectively.
  + Scikit-learn: Evaluation metrics—including accuracy, precision, recall, F1 scores, and confusion matrices.
  + PyTorch: Powered the development and training of deep neural networks, used to optimize learning rate schedules to improve efficiency.
  + YOLO (Ultralytics): Core detection model.
  + os and shutil: Utilities for managing file systems and organizing project directories effectively.
  + random: Application of controlled randomness, preserving consistency across multiple experiments.
  + scipy.stats: Statistical evaluation of results
  + itertools: efficiently generates all possible pairs, reducing redundant code

## Experimental Design and Model Validation

To ensure statistical validity and balanced training, the model was trained 10 separate times using a specific set of parameters. Three models were trained with image size 640, each with progressively increasing cumulative instances per classification group, namely 500, 1500, and 2500. Additionally, two models were trained using image size 1200—one solely increasing the image size and the other incorporating background augmentation to introduce variability into the dataset.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Name model | Instance threshold | image size | Training IoU | Training confidence | Validation IoU | Validation confidence |
| results\_500 | 500 | 640 | 0.7 | null | N/A | N/A |
| results\_1500 | 1500 | 640 | 0.7 | null | N/A | N/A |
| results\_2500 | 2500 | 640 | 0.7 | null | N/A | N/A |
| 2500\_1200 | 2500 | 1200 | 0.7 | null | N/A | N/A |
| 2500\_1200\_BG | 2500 | 1200 | 0.7 | null | N/A | N/A |
| val\_results\_500 | 500 | 640 | 0.7 | null | 0.15 | 0.35 |
| val\_results\_1500 | 1500 | 640 | 0.7 | null | 0.15 | 0.35 |
| val\_results\_2500 | 2500 | 640 | 0.7 | null | 0.15 | 0.35 |
| val\_2500\_1200 | 2500 | 1200 | 0.7 | null | 0.15 | 0.35 |
| val\_2500\_1200BG | 2500 | 1200 | 0.7 | null | 0.15 | 0.35 |

Table : Summary of Experimental Settings

All trained models underwent validation using different IoU and confidence threshold settings to ensure a rigorous assessment of model performance and classification accuracy across varying training configurations. Table 4 provides detailed information on the experimental settings used. The validation process examined image size conditions and augmentation effects, contributing to a comprehensive statistical analysis.

Due to the small sample size and uncertainty regarding normality, the Pairwise Mann-Whitney U Test was employed as a non-parametric statistical approach to compare performance assessing the following key metrics:

mAP50 (Mean Average Precision at IoU threshold 0.50)

mAP50-95 (Mean Average Precision across IoU thresholds from 0.50 to 0.95)

Recall (Ability to correctly detect relevant objects)

Precision (Proportion of correctly predicted objects among all detections)

Visual inspection and classification-specific calculations of prediction performance were conducted using data generated by the best-performing model, determined based on mAP50 and mAP50-95 scores. The model used was val\_2500\_1200BG.

### 4. Results

The experimental results reveal substantial performance differences between models trained with image size 1200 (2500\_1200, 2500\_1200\_BG) compared to those trained with image size 640 (results\_500, results\_1500, results\_2500). Moreover, notable improvements in detection accuracy are observed when adjusting IoU and confidence settings during validation runs. This is evident in the validation results of val\_results\_500, val\_results\_1500, val\_results\_2500, val\_2500\_1200, and val\_2500\_1200\_BG, which display enhanced detection capabilities compared to their respective original models.

During training, all models exhibited similar performance, as illustrated in Figure 6. Improvements were consistently observed in precision, mAP50, and mAP50-95 with still room for improvement, while recall reached its peak at epoch 30 (and for training with image size 640 around epoch 75). However, further training poses a risk of divergence between training and validation losses, including train/box\_loss vs. val/box\_loss, train/cls\_loss vs. val/cls\_loss, and train/dfl\_loss vs. val/dfl\_loss.

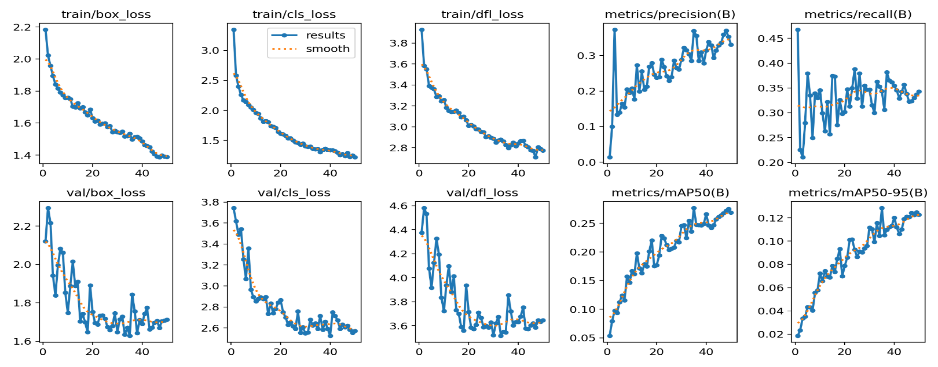


Figure : Training overview 2500\_1200\_BG

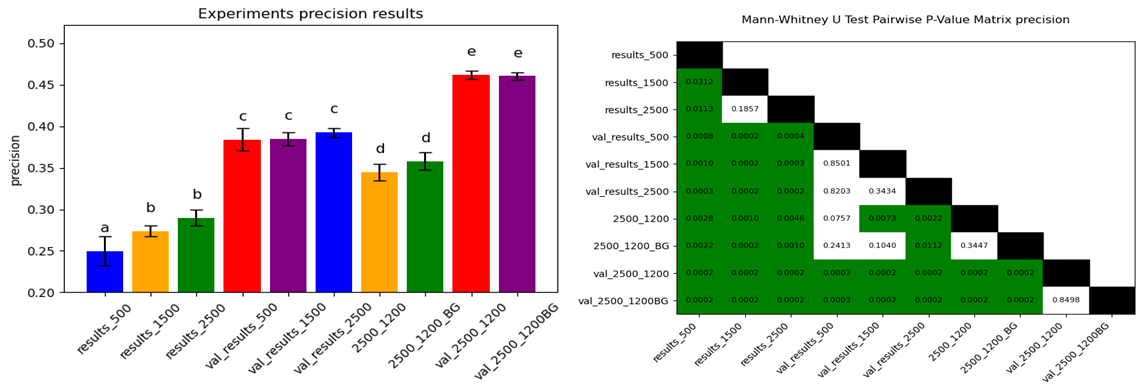
In terms of precision, val\_2500\_1200 and val\_2500\_1200BG achieved the highest scores (~0.45), surpassing models trained with image size 640 by at least 0.05 points. Additionally, statistical analysis using the Mann-Whitney U test confirms a significant difference with a p-value < 0.001 as seen in figure 7, highlighting the effectiveness of larger image sizes and refined validation settings in improving the model's precision performance.

Figure :Precision performance and statistical relevance

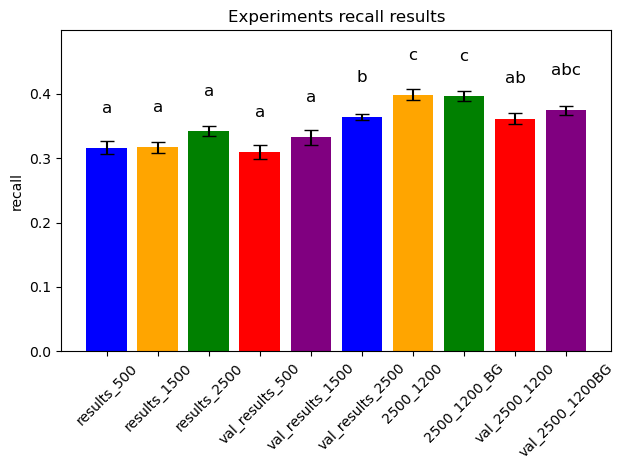
Regarding recall, the findings indicate that increasing the image size has a positive effect on this metric as seen in figure 8, resulting in an approximate 0.05-point improvement compared to models trained with smaller image sizes. The models achieving the highest recall scores include 2500\_1200, 2500\_1200\_BG, and val\_2500\_1200BG. However, adjustments to IoU and confidence settings in validation led to a decline in recall for models trained with an image size of 1200, highlighting the trade-offs associated with these parameter modifications.

Figure :Experiments recall performance

Experiments val\_2500\_1200, and val\_2500\_1200BG consistently achieve the highest accuracy across both mAP50 (~0.4) and mAP50-95 (~0.25). Moderate-performing models include val\_results\_500, val\_results\_1500, val\_results\_2500, 2500\_1200, and 2500\_1200\_BG, with mAP50 scores around 0.3-0.35 and mAP50-95 ranging between 0.12-0.2. Lower-scoring experiments include results\_500, results\_1500, and results\_2500 shown the lowest accuracy in both metrics, with mAP50 values between 0.18-0.22 and mAP50-95 between 0.07-0.1 as seen in figure 9.

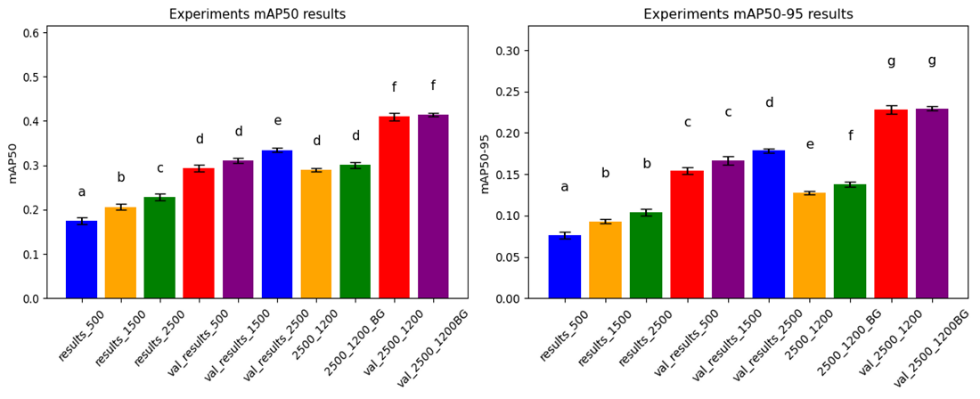


Figure :mAP50 and mAP50-95 performance

Table 5 presents the classification performance for images containing ORYSA for model val\_2500\_1200BG. Out of the total classifications, 79 were correctly identified, while 15 instances were missed. The most frequent misclassifications of ORYSA involved 1ECHG – Weed (100 incorrect or unmatched predictions) and ORYSN – Weed (24 instances). The primary issue in classification stems from unmatched predictions, referred to as "background" in Yolo. This occurs because Yolo often generates multiple bounding boxes for a single prediction, with only the best bounding box determining correctness, while the remaining predictions are categorized as background. Classification performance of the other classes for model val\_2500\_1200BG are added to appendix D.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | ORYSA | | 1ECHG - Weed | | CYPDI - Weed | ORYSN - Weed | | No Match |
| images | Correct | Unmatched Prediction | Incorrect | Unmatched Prediction | Unmatched Prediction | Incorrect | Unmatched Prediction | Unmatched Annotation |
| IMG\_20241021\_105827\_val | 5 | 6 |  | 1 |  |  |  |  |
| IMG\_20241021\_105838\_val | 5 | 10 | 18 | 20 |  |  | 1 | 1 |
| IMG\_20241022\_133012\_val | 7 | 6 |  |  |  |  |  | 1 |
| IMG\_20241022\_133021\_val | 6 | 15 | 17 | 19 |  |  |  |  |
| IMG\_20241029\_103205\_val | 4 | 5 | 5 | 7 |  | 11 | 7 | 7 |
| IMG\_20241029\_103212\_val | 8 | 6 |  |  | 1 |  |  |  |
| IMG\_20241029\_125217\_val | 2 | 3 |  |  |  | 1 | 2 |  |
| IMG\_20241101\_104638\_val | 1 | 2 |  |  |  |  |  | 2 |
| IMG\_20241101\_104642\_val | 1 | 1 |  |  |  |  |  | 2 |
| IMG\_20241101\_104648\_val | 9 | 12 | 3 | 5 |  |  |  | 1 |
| IMG\_20241104\_133857\_val | 14 | 13 | 1 | 3 |  |  |  |  |
| IMG\_20241104\_133901\_val | 2 | 6 | 1 | 1 |  |  |  |  |
| IMG\_20241217\_105511\_val | 2 | 3 |  |  |  |  | 1 | 1 |
| IMG\_20241217\_105516\_val | 10 | 10 |  |  |  |  |  |  |
| IMG\_20241217\_105547\_val | 3 | 3 |  |  |  |  | 1 |  |

Table : ORYSA classification performance

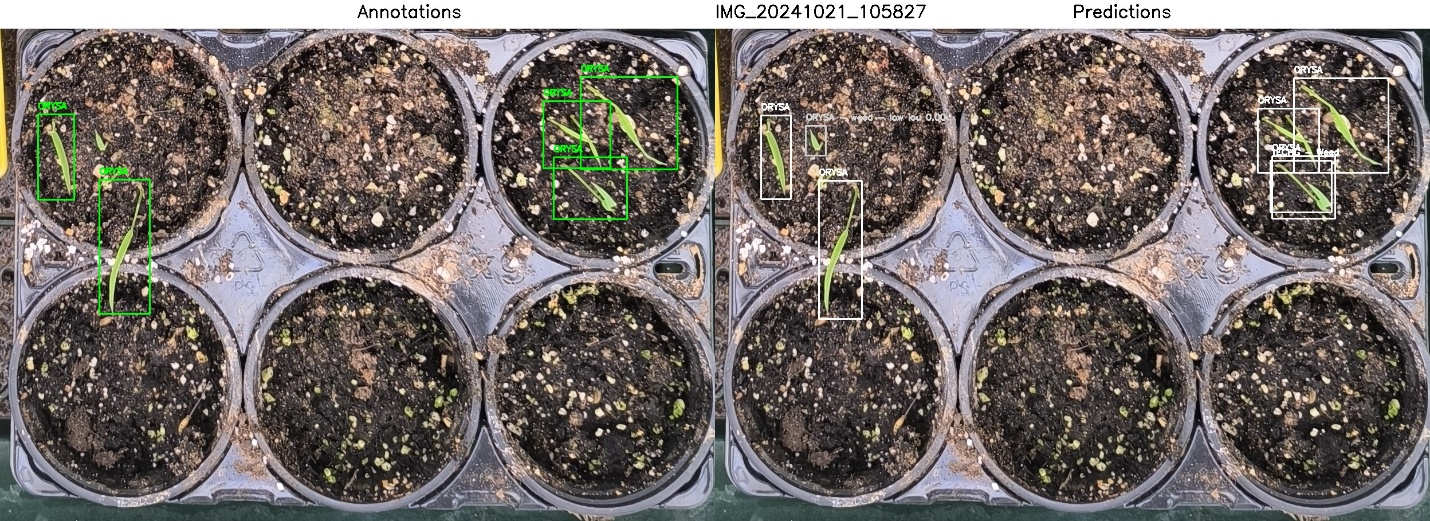


Figure : Annotation and prediction IMG\_20241021\_105827\_val

Figure 10 presents the initial annotated image on the left and the prediction for IMG\_20241021\_105827 on the right for model val\_2500\_1200BG, while Appendix E contains all validation images with annotations and predictions derived from val\_2500\_1200BG model. Original annotations are shown in green, with high-IoU predictions in white and low-IoU predictions in grey. All annotated instances were correctly predicted, although an unannotated area was classified as ORYSA, revealing an annotation error. Table 5 reports 5 correct predictions and 6 unmatched ones, and Table 6 further indicates that 5 of the 6 unmatched predictions closely resemble correct ones, as identified in its bottom row, where correct predictions are matched with similar unmatched predictions.

|  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Prediction\_ID** | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| **Category** | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | No Match | No Match | No Match | No Match | No Match | No Match | No Match |
| **x\_center** | 0.079 | 0.175 | 0.880 | 0.807 | 0.826 | 0.876 | 0.085 | 0.175 | 0.802 | 0.823 | 0.141 | 0.820 |
| **y\_center** | 0.261 | 0.442 | 0.192 | 0.215 | 0.325 | 0.197 | 0.261 | 0.450 | 0.228 | 0.322 | 0.228 | 0.320 |
| **Width** | 0.051 | 0.071 | 0.135 | 0.095 | 0.103 | 0.132 | 0.042 | 0.061 | 0.086 | 0.088 | 0.029 | 0.086 |
| **Height** | 0.173 | 0.272 | 0.189 | 0.137 | 0.126 | 0.192 | 0.170 | 0.280 | 0.131 | 0.128 | 0.060 | 0.105 |
| **Predicted Category** | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | ORYSA | 1ECHG - Weed |
| **Match Status** | Correct | Correct | Correct | Correct | Correct | Unmatched Prediction | Unmatched Prediction | Unmatched Prediction | Unmatched Prediction | Unmatched Prediction | Unmatched Prediction | Unmatched Prediction |
| **Possible Match** |  |  |  |  |  | 6 -> 3 | 7 -> 1 | 8 -> 2 | 9 -> 4 | 10 -> 5 |  | 12 -> 5 |

Table : Similarity check Yolo predictions (IMG\_20241021\_105827\_val)

### 6. Discussion and conclusion

## Classification Performance & Challenges

Classification accuracy is largely influenced by image resolution, validation settings, and annotation methods. Models trained using image size 1200 demonstrate superior performance compared to image size 640, achieving higher mAP50 and mAP50-95 scores. However, adjustments in validation parameters introduce a trade-off between precision and recall, requiring further optimization to achieve a balanced improvement in accuracy.

Visual observations reveal that classification performance declines when large overlapping groups are present. When multiple objects significantly overlap in an image, the model struggles to differentiate between them, leading to misclassifications or background predictions. This is particularly evident in cases where ORYSA and ORYSN are misclassified as 1ECHG – Weed, indicating that highly similar classes affect prediction reliability.

## Annotation & Prediction Discrepancies

The training dataset underwent a high degree of augmentation, significantly increasing image variations for model learning. While augmentation enhances diversity and robustness, it also introduces vulnerability to annotation errors. When augmentation is applied extensively, small inaccuracies in initial annotations can be amplified, leading to misaligned bounding boxes and classification inconsistencies. This effect is particularly noticeable when objects overlap or when annotation patterns diverge slightly from their original positioning.

Figure 10 presents a clear example of this issue, where a possibly missed annotation was predicted as ORYSN - Weed, revealing likely an annotation error. Such inconsistencies highlight the need for careful annotation adjustments to minimize unintended inaccuracies that may distort model predictions.

Additionally, instead of relying primarily on augmentation, increasing the number of original images can enhance specificity, while higher-resolution images can help the model detect finer details. A more diverse dataset with naturally varied high-resolution images could improve the model’s ability to differentiate between closely related classes, reducing dependence on artificial diversity and minimizing the amplification of annotation errors.

The results for CYPDI – Weed reveal differences between annotation and prediction. This inconsistency occurs because CYPDI is annotated as groups rather than individually, due to the impracticality of marking individual instances of this species. Consequently, predictions contain multiple bounding boxes, which may differ in placement but still accurately reflect the expected class. These discrepancies highlight limitations in annotation methodology rather than classification errors.

For 1ECHG – Weed, it emerges as the most frequent misclassification for ORYSA and ORYSN. Additionally, for 1ECHG classifications, ORYSN is the most common misclassification, followed by ORYSA, demonstrating a recurring pattern in prediction errors. The similarity in features among these classes likely contributes to confusion, and refining annotation techniques, increasing image resolution, or improving feature extraction could help mitigate these errors.

A potential solution to enhance the classification performance of 1ECHG – Weed is to introduce a size-based bias, where 1ECHG is given a larger annotation size than other categories. This slight bias could aid the model in distinguishing 1ECHG from ORYSA and ORYSN, thereby reducing misclassification errors.

Background predictions in Yolo are another factor affecting classification accuracy. Since Yolo can generate multiple bounding boxes per prediction, it retains only the best match as correct classification, while the remaining bounding boxes are labeled as background. As a result, certain misclassifications are categorized as background, further affecting accuracy measurements. Applying post-processing techniques to refine the model’s output by eliminating background-labeled predictions that closely resemble correct classifications can improve accuracy assessments.

## Key Observations

* Large overlapping groups negatively impact classification accuracy, increasing misclassification rates.
* CYPDI annotation constraints lead to different bounding box patterns compared to predictions, but the classifications remain accurate.
* 1ECHG is the most frequent misclassification for ORYSA and ORYSN, while ORYSN is the most common misclassification for 1ECHG, followed by ORYSA.
* Annotation inconsistencies contribute to prediction errors, which could be compounded by the high degree of image augmentation used.

## Practical Applications for Farmers

The developed model has significant practical implications for precision agriculture, particularly in weed detection and crop classification. By integrating this system into automated farming solutions, farmers can enhance weed management, reduce manual labor, and improve crop health.

* Improved Weed Control: The model accurately identifies weeds such as 1ECHG – Weed and ORYSN – Weed, facilitating precise herbicide application while minimizing unintended exposure to crops.
* Autonomous Farming Solutions: The model can be integrated with robotic weed removal systems, enabling automated weed extraction while ensuring essential crops remain unaffected.
* IoT & Smart Farming Integration: Connecting the model with drone-based monitoring systems or IoT sensors enables real-time weed detection and automated responses, making large-scale agriculture more efficient.

### 7. References

Araújo, S.O., Peres, R.S., Ramalho, J.C., Lidon, F. and Barata, J., 2023. Machine learning applications in agriculture: current trends, challenges, and future perspectives. \*Agronomy\*, 13(12), p.2976.

Botero-Valencia, J., García-Pineda, V., Valencia-Arias, A., Valencia, J., Reyes-Vera, E., Mejia-Herrera, M. and Hernández-García, R., 2025. Machine Learning in Sustainable Agriculture: Systematic Review and Research Perspectives. \*Agriculture\*, 15(4), p.377.

Chauhan, B.S., 2012. Weed ecology and weed management strategies for dry-seeded rice in Asia. \*Weed Technology\*, 26(1), pp.1-13.

European Commission, 2025. Rice Production Dashboard. Available at: https://agridata.ec.europa.eu/extensions/DashboardRice/RiceProduction.html [Accessed 5 Apr. 2025]

Godar, A.S. and Norsworthy, J.K., 2023. *Echinochloa* in mid-southern US and California rice: What is known and what are the knowledge gaps?. \*Weed Technology\*, 37(4), pp.452-469.

Human Signal. (n.d.). LabelImg. Available at: <https://github.com/HumanSignal/labelImg/releases> (Accessed: 5 April 2025).

Nadir, S., Xiong, H.B., Zhu, Q., Zhang, X.L., Xu, H.Y., Li, J., Dongchen, W., Henry, D., Guo, X.Q., Khan, S. and Suh, H.S., 2017. Weedy rice in sustainable rice production. A review. Agronomy for Sustainable Development, 37, pp.1-14.

Stroppiana, D., Villa, P., Sona, G., Ronchetti, G., Candiani, G., Pepe, M., Busetto, L., Migliazzi, M. and Boschetti, M., 2018. Early season weed mapping in rice crops using multi-spectral UAV data. \*International Journal of Remote Sensing\*, 39(15-16), pp.5432-5452.

Sustainable EU Rice Project, 2025. European Rice Overview. [online] Available at: <https://www.sustainableeurice.eu/european-rice/> [Accessed 5 Apr. 2025].

Tian, Z., Shen, G., Yuan, G., Song, K., Lu, J. and Da, L., 2020. Effects of *Echinochloa crusgalli* and *Cyperus difformis* on yield and eco-economic thresholds of rice. \*Journal of Cleaner Production\*, 259, p.120807.

Vulchi, R., Guan, T., Clark, T. and Brim-DeForest, W., 2024. Echinochloa spp response to preemergence and postemergence herbicides in California rice (Oryza sativa L.). Frontiers in Agronomy, 6, p.1349008.

Wu, Z., Chen, Y., Zhao, B., Kang, X. and Ding, Y., 2021. Review of weed detection methods based on computer vision. \*Sensors\*, 21(11), p.3647.

Yuan, G., Gao, Y., Fang, J., Shen, G. and Tian, Z., 2025. Environmental Influences on Seed Germination and Seedling Emergence in Four *Echinochloa* Taxa. \*Agronomy\*, 15(2).

Ziska, L.H., Gealy, D.R., Burgos, N., Caicedo, A.L., Gressel, J., Lawton-Rauh, A.L., Avila, L.A., Theisen, G., Norsworthy, J., Ferrero, A. and Vidotto, F., 2015. Weedy (red) rice: an emerging constraint to global rice production. \*Advances in Agronomy\*, 129, pp.181-228.

### Appendices

### Appendix A: Images included in training detection model

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Filename | Width | Height | Resolution (DPI) | BrightnessValue | LightSource | SensingMethod | ExposureProgram | Flash | ISOSpeedRatings |
| 1730361965076.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730361965097.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364827176.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364827194.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364858393.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364858405.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364906816.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364906832.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364930953.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364954263.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| 1730364954270.jpg | 3072 | 4080 | (72.0, 72.0) | 0 | 21 | 0 | 0 | no data | no data |
| IMG\_20241021\_105827.jpg | 2296 | 4080 | (72.0, 72.0) | 2.6 | 255 | no data | 2 | 16 | 500 |
| IMG\_20241021\_105838.jpg | 2296 | 4080 | (72.0, 72.0) | 3.1 | 255 | no data | 2 | 16 | 320 |
| IMG\_20241021\_105851.jpg | 2296 | 4080 | (72.0, 72.0) | 2.3 | 255 | no data | 2 | 16 | 500 |
| IMG\_20241021\_105900.jpg | 2296 | 4080 | (72.0, 72.0) | 2.8 | 255 | no data | 2 | 16 | 500 |
| IMG\_20241021\_105912.jpg | 2296 | 4080 | (72.0, 72.0) | 3.7 | 255 | no data | 2 | 16 | 250 |
| IMG\_20241021\_105916.jpg | 2296 | 4080 | (72.0, 72.0) | 2.9 | 255 | no data | 2 | 16 | 400 |
| IMG\_20241021\_105930.jpg | 2296 | 4080 | (72.0, 72.0) | 3.4 | 255 | no data | 2 | 16 | 400 |
| IMG\_20241021\_110110.jpg | 2296 | 4080 | (72.0, 72.0) | 3.4 | 255 | no data | 2 | 16 | 500 |
| IMG\_20241021\_110623.jpg | 2296 | 4080 | (72.0, 72.0) | 5 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241022\_132953.jpg | 2296 | 4080 | (72.0, 72.0) | 5.1 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241022\_133002.jpg | 2296 | 4080 | (72.0, 72.0) | 5.1 | 255 | no data | 2 | 16 | 100 |
| IMG\_20241022\_133012.jpg | 2296 | 4080 | (72.0, 72.0) | 5 | 255 | no data | 2 | 16 | 80 |
| IMG\_20241022\_133021.jpg | 2296 | 4080 | (72.0, 72.0) | 5.1 | 255 | no data | 2 | 16 | 80 |
| IMG\_20241022\_133031.jpg | 2296 | 4080 | (72.0, 72.0) | 4.8 | 255 | no data | 2 | 16 | 100 |
| IMG\_20241022\_133040.jpg | 2296 | 4080 | (72.0, 72.0) | 4.9 | 255 | no data | 2 | 16 | 100 |
| IMG\_20241022\_133052.jpg | 2296 | 4080 | (72.0, 72.0) | 4.8 | 255 | no data | 2 | 16 | 100 |
| IMG\_20241022\_133100.jpg | 2296 | 4080 | (72.0, 72.0) | 4.8 | 255 | no data | 2 | 16 | 100 |
| IMG\_20241029\_103126.jpg | 2296 | 4080 | (72.0, 72.0) | 4.3 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_103133.jpg | 2296 | 4080 | (72.0, 72.0) | 4.4 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_103150.jpg | 2296 | 4080 | (72.0, 72.0) | 4.2 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_103157.jpg | 2296 | 4080 | (72.0, 72.0) | 4.1 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_103205.jpg | 4080 | 2296 | (72.0, 72.0) | 4.2 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_103212.jpg | 2296 | 4080 | (72.0, 72.0) | 4.1 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_124651.jpg | 2296 | 4080 | (72.0, 72.0) | 4.7 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_125158.jpg | 4080 | 2296 | (72.0, 72.0) | 4.6 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_125206.jpg | 4080 | 2296 | (72.0, 72.0) | 4.4 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_125217.jpg | 4080 | 2296 | (72.0, 72.0) | 4.2 | 255 | no data | 2 | 16 | 125 |
| IMG\_20241029\_125222.jpg | 4080 | 2296 | (72.0, 72.0) | 4.5 | 255 | no data | 2 | 16 | 160 |
| IMG\_20241029\_125226.jpg | 4080 | 2296 | (72.0, 72.0) | 4.8 | 255 | no data | 2 | 16 | 160 |
| IMG\_20241101\_104154.jpg | 2296 | 4080 | (72.0, 72.0) | 4.7 | 255 | no data | 2 | 24 | 160 |
| IMG\_20241101\_104253.jpg | 2296 | 4080 | (72.0, 72.0) | 4.8 | 255 | no data | 2 | 24 | 160 |
| IMG\_20241101\_104432.jpg | 4080 | 2296 | (72.0, 72.0) | 4.3 | 255 | no data | 2 | 24 | 200 |
| IMG\_20241101\_104528.jpg | 2296 | 4080 | (72.0, 72.0) | 4.2 | 255 | no data | 2 | 24 | 200 |
| IMG\_20241101\_104638.jpg | 2296 | 4080 | (72.0, 72.0) | 3.8 | 255 | no data | 2 | 24 | 160 |
| IMG\_20241101\_104642.jpg | 2296 | 4080 | (72.0, 72.0) | 3.7 | 255 | no data | 2 | 24 | 160 |
| IMG\_20241101\_104648.jpg | 2296 | 4080 | (72.0, 72.0) | 3.8 | 255 | no data | 2 | 24 | 160 |
| IMG\_20241101\_104701.jpg | 2296 | 4080 | (72.0, 72.0) | 4 | 255 | no data | 2 | 24 | 125 |
| IMG\_20241101\_104717.jpg | 2296 | 4080 | (72.0, 72.0) | 3.9 | 255 | no data | 2 | 24 | 125 |
| IMG\_20241101\_104743.jpg | 2296 | 4080 | (72.0, 72.0) | 3.8 | 255 | no data | 2 | 24 | 125 |
| IMG\_20241101\_104752.jpg | 2296 | 4080 | (72.0, 72.0) | 4.3 | 255 | no data | 2 | 24 | 125 |
| IMG\_20241101\_104804.jpg | 2296 | 4080 | (72.0, 72.0) | 4.4 | 255 | no data | 2 | 24 | 100 |
| IMG\_20241101\_104806.jpg | 2296 | 4080 | (72.0, 72.0) | 4.3 | 255 | no data | 2 | 24 | 100 |
| IMG\_20241101\_104810.jpg | 2296 | 4080 | (72.0, 72.0) | 4.4 | 255 | no data | 2 | 24 | 125 |
| IMG\_20241104\_133835.jpg | 2296 | 4080 | (72.0, 72.0) | 2.2 | 255 | no data | 2 | 24 | 500 |
| IMG\_20241104\_133840.jpg | 2296 | 4080 | (72.0, 72.0) | 2.3 | 255 | no data | 2 | 24 | 500 |
| IMG\_20241104\_133846.jpg | 2296 | 4080 | (72.0, 72.0) | 2.1 | 255 | no data | 2 | 24 | 500 |
| IMG\_20241104\_133851.jpg | 2296 | 4080 | (72.0, 72.0) | 1.8 | 255 | no data | 2 | 24 | 500 |
| IMG\_20241104\_133857.jpg | 2296 | 4080 | (72.0, 72.0) | 1.9 | 255 | no data | 2 | 24 | 640 |
| IMG\_20241104\_133901.jpg | 2296 | 4080 | (72.0, 72.0) | 1.9 | 255 | no data | 2 | 24 | 640 |
| IMG\_20241104\_134050.jpg | 2296 | 4080 | (72.0, 72.0) | 2.3 | 255 | no data | 2 | 24 | 800 |
| IMG\_20241104\_134131.jpg | 2296 | 4080 | (72.0, 72.0) | 3.1 | 255 | no data | 2 | 24 | 500 |
| IMG\_20241104\_134344.jpg | 2296 | 4080 | (72.0, 72.0) | 2.1 | 255 | no data | 2 | 24 | 800 |
| IMG\_20241217\_103801.jpg | 2296 | 4080 | (72.0, 72.0) | 3.5 | 255 | no data | 2 | 16 | 160 |
| IMG\_20241217\_103806.jpg | 2296 | 4080 | (72.0, 72.0) | 3.9 | 255 | no data | 2 | 16 | 200 |
| IMG\_20241217\_104107.jpg | 2296 | 4080 | (72.0, 72.0) | 3.9 | 255 | no data | 2 | 16 | 200 |
| IMG\_20241217\_104116.jpg | 2296 | 4080 | (72.0, 72.0) | 3.9 | 255 | no data | 2 | 16 | 200 |
| IMG\_20241217\_105511.jpg | 2296 | 4080 | (72.0, 72.0) | 4.2 | 255 | no data | 2 | 16 | 160 |
| IMG\_20241217\_105516.jpg | 2296 | 4080 | (72.0, 72.0) | 4.1 | 255 | no data | 2 | 16 | 160 |
| IMG\_20241217\_105547.jpg | 2296 | 4080 | (72.0, 72.0) | 4.1 | 255 | no data | 2 | 16 | 160 |

### Appendix B: Model settings used for image size 640

|  |  |  |
| --- | --- | --- |
| epochs: 150 (default: 100) | conf: null (default: null) | momentum: 0.937 (default: 0.937) |
| time: null (default: null) | iou: 0.7 (default: 0.7) | weight\_decay: 0.0005 (default: 0.0005) |
| patience: 100 (default: 50) | max\_det: 300 (default: 300) | warmup\_epochs: 3.0 (default: 3.0) |
| batch: 16 (default: 16) | half: false (default: false) | warmup\_momentum: 0.8 (default: 0.8) |
| imgsz: 640 (default: 640) | dnn: false (default: false) | warmup\_bias\_lr: 0.1 (default: 0.1) |
| save: true (default: true) | plots: true (default: true) | box: 7.5 (default: 7.5) |
| save\_period: -1 (default: -1) | source: null (default: null) | cls: 0.5 (default: 0.5) |
| cache: false (default: false) | vid\_stride: 1 (default: 1) | dfl: **3** (default: 4) |
| device: null (default: null) | stream\_buffer: false (default: false) | pose: 12.0 (default: 12.0) |
| workers: 8 (default: 8) | visualize: false (default: false) | kobj: 1.0 (default: 1.0) |
| project: C:/**~**/ (default: path/to/project/) | augment: false (default: false) | nbs: 64 (default: 64) |
| name: train2 (default: exp) | agnostic\_nms: false (default: false) | hsv\_h: 0.015 (default: 0.015) |
| exist\_ok: false (default: false) | classes: null (default: null) | hsv\_s: 0.7 (default: 0.7) |
| pretrained: true (default: true) | retina\_masks: false (default: false) | hsv\_v: 0.4 (default: 0.4) |
| optimizer: Adam (default: SGD) | embed: null (default: null) | degrees: 0.0 (default: 0.0) |
| verbose: true (default: true) | show: false (default: false) | translate: 0.1 (default: 0.1) |
| seed: 0 (default: 0) | save\_frames: false (default: false) | scale: 0.5 (default: 0.5) |
| deterministic: true (default: true) | save\_txt: false (default: false) | shear: 0.0 (default: 0.0) |
| single\_cls: false (default: false) | save\_conf: false (default: false) | perspective: 0.0 (default: 0.0) |
| rect: false (default: false) | save\_crop: false (default: false) | flipud: 0.0 (default: 0.0) |
| cos\_lr: false (default: false) | show\_labels: true (default: true) | fliplr: 0.5 (default: 0.5) |
| close\_mosaic: 10 (default: 10) | show\_conf: true (default: true) | bgr: 0.0 (default: 0.0) |
| resume: false (default: false) | show\_boxes: true (default: true) | mosaic: 1.0 (default: 1.0) |
| amp: true (default: true) | line\_width: null (default: null) | mixup: 0.0 (default: 0.0) |
| fraction: 1.0 (default: 1.0) | format: torchscript (default: torchscript) | copy\_paste: 0.0 (default: 0.0) |
| profile: false (default: false) | keras: false (default: false) | copy\_paste\_mode: flip (default: flip) |
| freeze: null (default: null) | optimize: false (default: false) | auto\_augment: randaugment (default: randaugment) |
| multi\_scale: false (default: false) | int8: false (default: false) | erasing: 0.4 (default: 0.4) |
| overlap\_mask: true (default: true) | dynamic: false (default: false) | crop\_fraction: 1.0 (default: 1.0) |
| mask\_ratio: 4 (default: 4) | simplify: true (default: true) | cfg: null (default: null) |
| dropout: 0.05 (default: 0.05) | opset: null (default: null) | tracker: botsort.yaml (default: botsort.yaml) |

### Appendix C: Model settings used for image size 1200

|  |  |  |
| --- | --- | --- |
| task: detect (default: detect) | val: true (default: true) | workspace: null (default: null) |
| mode: train (default: train) | split: val (default: val) | nms: false (default: false) |
| model: yolov8n.pt (default: yolov8n.pt) | save\_json: false (default: false) | lr0: 0.0003 (default: 0.01) |
| data: C:/~/data.yaml (default: path/to/data.yaml) | save\_hybrid: false (default: false) | lrf: 0.01 (default: 0.01) |
| epochs: 50 (default: 100) | conf: null (default: null) | momentum: 0.937 (default: 0.937) |
| time: null (default: null) | iou: 0.7 (default: 0.7) | weight\_decay: 0.0005 (default: 0.0005) |
| patience: 100 (default: 50) | max\_det: 300 (default: 300) | warmup\_epochs: 3.0 (default: 3.0) |
| batch: 9 (default: 16) | half: false (default: false) | warmup\_momentum: 0.8 (default: 0.8) |
| imgsz: 1200 (default: 640) | dnn: false (default: false) | warmup\_bias\_lr: 0.1 (default: 0.1) |
| save: true (default: true) | plots: true (default: true) | box: 7.5 (default: 7.5) |
| save\_period: -1 (default: -1) | source: null (default: null) | cls: 0.5 (default: 0.5) |
| cache: false (default: false) | vid\_stride: 1 (default: 1) | dfl: 3 (default: 4) |
| device: null (default: null) | stream\_buffer: false (default: false) | pose: 12.0 (default: 12.0) |
| workers: 8 (default: 8) | visualize: false (default: false) | kobj: 1.0 (default: 1.0) |
| project: C:/~/yolo\_data/ (default: path/to/project/) | augment: false (default: false) | nbs: 64 (default: 64) |
| name: train9 (default: exp) | agnostic\_nms: false (default: false) | hsv\_h: 0.015 (default: 0.015) |
| exist\_ok: false (default: false) | classes: null (default: null) | hsv\_s: 0.7 (default: 0.7) |
| pretrained: true (default: true) | retina\_masks: false (default: false) | hsv\_v: 0.4 (default: 0.4) |
| optimizer: Adam (default: SGD) | embed: null (default: null) | degrees: 0.0 (default: 0.0) |
| verbose: true (default: true) | show: false (default: false) | translate: 0.1 (default: 0.1) |
| seed: 0 (default: 0) | save\_frames: false (default: false) | scale: 0.5 (default: 0.5) |
| deterministic: true (default: true) | save\_txt: false (default: false) | shear: 0.0 (default: 0.0) |
| single\_cls: false (default: false) | save\_conf: false (default: false) | perspective: 0.0 (default: 0.0) |
| rect: false (default: false) | save\_crop: false (default: false) | flipud: 0.0 (default: 0.0) |
| cos\_lr: false (default: false) | show\_labels: true (default: true) | fliplr: 0.5 (default: 0.5) |
| close\_mosaic: 10 (default: 10) | show\_conf: true (default: true) | bgr: 0.0 (default: 0.0) |
| resume: false (default: false) | show\_boxes: true (default: true) | mosaic: 1.0 (default: 1.0) |
| amp: true (default: true) | line\_width: null (default: null) | mixup: 0.0 (default: 0.0) |
| fraction: 1.0 (default: 1.0) | format: torchscript (default: torchscript) | copy\_paste: 0.0 (default: 0.0) |
| profile: false (default: false) | keras: false (default: false) | copy\_paste\_mode: flip (default: flip) |
| freeze: null (default: null) | optimize: false (default: false) | auto\_augment: randaugment (default: randaugment) |
| multi\_scale: false (default: false) | int8: false (default: false) | erasing: 0.4 (default: 0.4) |
| overlap\_mask: true (default: true) | dynamic: false (default: false) | crop\_fraction: 1.0 (default: 1.0) |
| mask\_ratio: 4 (default: 4) | simplify: true (default: true) | cfg: null (default: null) |
| dropout: 0.15 (default: 0.05) | opset: null (default: null) | tracker: botsort.yaml (default: botsort.yaml) |

### Appendix D: Classification performance

ORYSN classification performance

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | ORYSN - Weed | | 1ECHG - Weed | | ORYSA | | No Match |
| Row Labels | Correct | Unmatched Prediction | Incorrect | Unmatched Prediction | Incorrect | Unmatched Prediction | Unmatched Annotation |
| IMG\_20241021\_105930\_val.txt | 1 | 1 | 2 | 2 |  |  |  |
| IMG\_20241021\_110110\_val.txt | 6 | 6 |  | 1 |  |  | 1 |
| IMG\_20241021\_110623\_val.txt | 8 | 10 |  |  |  | 1 |  |
| IMG\_20241022\_132953\_val.txt | 7 | 5 | 3 | 2 |  |  | 1 |
| IMG\_20241022\_133002\_val.txt | 7 | 8 |  |  |  |  |  |
| IMG\_20241029\_125222\_val.txt | 6 | 4 |  |  |  | 1 |  |
| IMG\_20241029\_125226\_val.txt | 1 | 3 |  |  |  |  | 2 |
| IMG\_20241101\_104154\_val.txt | 1 | 2 |  |  | 1 | 1 |  |
| IMG\_20241101\_104253\_val.txt | 1 | 2 | 1 | 1 |  |  |  |
| IMG\_20241101\_104432\_val.txt | 2 | 3 |  |  |  |  | 2 |
| IMG\_20241101\_104528\_val.txt | 3 | 3 |  |  |  |  |  |
| IMG\_20241104\_134050\_val.txt | 3 | 3 |  | 1 | 2 | 1 |  |
| IMG\_20241104\_134131\_val.txt | 7 | 6 |  |  |  | 1 | 1 |

CYPDI classification performance

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | CYPDI - Weed | | 1ECHG - Weed | | No Match |
| Row Labels | Correct | Unmatched Prediction | Incorrect | Unmatched Prediction | Unmatched Annotation |
| IMG\_20241021\_105912\_val.txt | 1 | 2 |  |  |  |
| IMG\_20241021\_105916\_val.txt | 3 | 2 |  |  |  |
| IMG\_20241022\_133052\_val.txt | 5 | 7 |  |  | 1 |
| IMG\_20241022\_133100\_val.txt | 3 | 3 |  |  | 1 |
| IMG\_20241029\_103126\_val.txt | 3 | 8 |  |  | 2 |
| IMG\_20241029\_103133\_val.txt | 10 | 17 |  |  | 1 |
| IMG\_20241101\_104804\_val.txt | 9 | 10 |  |  | 6 |
| IMG\_20241101\_104810\_val.txt | 4 | 9 |  |  | 6 |
| IMG\_20241104\_133835\_val.txt | 4 | 8 |  |  | 5 |
| IMG\_20241104\_133840\_val.txt | 8 | 12 |  |  | 3 |
| IMG\_20241217\_103801\_val.txt | 4 | 3 | 2 | 4 | 2 |
| IMG\_20241217\_103806\_val.txt | 3 | 3 | 4 | 7 | 3 |

1ECHG classification performance

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1ECHG - Weed | | CYPDI - Weed | ORYSA | | ORYSN - Weed | | No Match |
| Row Labels | Correct | Unmatched Prediction | Unmatched Prediction | Incorrect | Unmatched Prediction | Incorrect | Unmatched Prediction | Unmatched Annotation |
| IMG\_20241021\_105851\_val.txt | 14 | 11 |  |  |  |  |  | 2 |
| IMG\_20241021\_105900\_val.txt | 21 | 21 | 1 | 1 | 1 |  |  |  |
| IMG\_20241022\_133031\_val.txt | 19 | 17 |  |  |  |  |  |  |
| IMG\_20241022\_133040\_val.txt | 23 | 22 |  | 5 | 4 |  |  | 5 |
| IMG\_20241029\_103150\_val.txt | 6 | 5 |  |  |  | 12 | 5 | 12 |
| IMG\_20241029\_103157\_val.txt | 19 | 17 |  |  |  | 2 | 1 | 2 |
| IMG\_20241029\_124651\_val.txt | 12 | 12 |  |  |  |  |  |  |
| IMG\_20241029\_125158\_val.txt | 12 | 13 |  |  |  |  |  |  |
| IMG\_20241029\_125206\_val.txt | 9 | 10 |  |  |  | 3 | 2 |  |
| IMG\_20241101\_104701\_val.txt | 13 | 12 |  |  | 1 |  | 1 | 1 |
| IMG\_20241101\_104717\_val.txt | 12 | 12 |  |  |  |  |  | 1 |
| IMG\_20241101\_104743\_val.txt | 13 | 12 |  |  |  |  | 1 | 1 |
| IMG\_20241101\_104752\_val.txt | 12 | 9 |  |  | 1 |  |  |  |
| IMG\_20241104\_133846\_val.txt | 8 | 8 |  |  |  | 3 | 2 | 1 |
| IMG\_20241104\_133851\_val.txt | 10 | 12 |  |  | 2 | 1 | 1 | 1 |
| IMG\_20241217\_104107\_val.txt | 4 | 5 |  |  |  | 2 | 3 |  |
| IMG\_20241217\_104116\_val.txt | 1 | 1 |  | 1 | 3 |  | 1 | 1 |