Literature Review

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
18.	Jiawei Zhao, Guangzhao Tian, Chang Qiu, et al.	2022	Weed Detection in Potato Fields Based on Improved YOLOv4: Optimal Speed and Accuracy of Weed Detection in Potato Fields	**PROBLEM:** Standard object detection models like YOLOv4 often struggle to balance high accuracy with the real-time speed needed for on-field hardware. The authors developed a highly optimized model, **MC-YOLOv4**, by improving the standard YOLOv4. They replaced the heavy backbone with a lightweight **MobileNetV3**, introduced an attention mechanism (**CBAM**) to focus on important features, and used other advanced techniques to create a fast and efficient model for drawing **bounding boxes** around weeds in potato fields.	The study created a novel, high-performance model (MC-YOLOv4) that significantly outperformed standard YOLOv4 and other common detectors. Their model achieved an extremely high detection accuracy (**98.52% mAP**) while being significantly smaller and faster than the original YOLOv4. This demonstrated the value of lightweight architectures and attention mechanisms for real-time agricultural applications.	However, the study was performed under controlled laboratory-like conditions with clean backgrounds. The model did not account for field challenges such as soil interference, shadows, or mixed pixels, nor was scalability tested in real-world outdoor environments.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
19.	Ignazio Gallo, Anwar Ur Rehman,Ramin Heidarian Dekhordi, et al	2023	Deep Object Detection of Crop Weeds: Performance of YOLOv7 on a Real Case Dataset from UAV Images	**PROBLEM:** A lack of large, case-specific weed datasets from UAVs hinders the development of robust deep learning models for real-world agriculture. The authors addressed this by first creating a new, public dataset of chicory crops and weeds (CP dataset). They then benchmarked the performance of the **YOLOv7 object detection model**, a state-of-the-art algorithm at the time, to locate and draw **bounding boxes** around weeds.	The study made two key contributions: (1) It introduced a new, publicly available UAV weed dataset, which is a valuable resource for the research community. (2) It established a new state-of-the-art performance benchmark by demonstrating that **YOLOv7 outperforms previous YOLO versions** for weed detection from aerial imagery.	Even this cutting-edge 2023 study focuses exclusively on **object detection**. Its most precise output is a **bounding box**, which is an approximation of a weed's location. The work does not venture into instance segmentation, thereby lacking the pixel-level accuracy needed to determine the exact shape, size, and biomass of weeds for the most advanced precision agriculture automation.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
3.	Hao-Ran Qu and Wen-Hao Su	2024	Deep Learning-Based Weed–Crop Recognition for Smart Agricultural Equipment: A Review	**PROBLEM: Although many computer vision methods for weed detection exist, there is still a lack of systems that work in real time and can be deployed directly on farm machinery such as smart tractors or drones. This survey paper consolidates research on deep learning methods for crop–weed recognition to highlight where progress has been made and where gaps still exist.** The authors reviewed a wide range of recent studies, comparing object detection approaches (like YOLO and Faster R-CNN) with semantic segmentation methods (like U-Net and DeepLab). They also analyzed datasets used, model accuracies, and hardware requirements for deployment.	The survey concluded that YOLOv5 and YOLOv7 were particularly strong candidates for real-time weed detection due to their speed and accuracy balance, while U-Net and DeepLab models excelled at pixel-level segmentation but at higher computational cost. Across the board, deep learning outperformed traditional machine learning, especially in complex and cluttered field environments.	However, the majority of these deep learning approaches were developed with RGB or multispectral data, with hyperspectral deep learning applications still in their infancy. Another major gap was the lack of large, publicly available hyperspectral datasets for weeds and crops, which restricts the advancement of HSI-driven deep learning.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
4.	Inbal Ronay, Ran Nisim Lati and Fadi Kizel	2024	Weed Species Identification: Acquisition, Feature Analysis, and Evaluation of a Hyperspectral and RGB Dataset with Labeled Data	**PROBLEM: One of the biggest barriers in weed detection research is the absence of large, open-access datasets that allow researchers to benchmark models. This paper directly addresses that gap by releasing a new weed dataset.** The authors compiled a dataset consisting of 30 hyperspectral and RGB images of common weed species. These were annotated at the pixel level for weed–crop separation. Baseline machine learning and deep learning methods were run on the dataset to provide reference benchmarks.	The dataset proved useful for testing both spectral-only and spectral–spatial models. While classical machine learning methods were effective for species-level classification, deep learning approaches were much stronger for pixel-level segmentation. Importantly, the dataset is openly available, providing the research community with a shared benchmark for future work.	However, the dataset was created under controlled conditions and is relatively small, covering only a limited number of weed species and growth stages. It does not capture the complexity of UAV-scale imagery in real-world field environments.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
5.	Tao Liu, Yuanyuan Zhao, Li et al.	2024	Harnessing UAVs and deep learning for accurate grass weed detection in wheat fields: a study on biomass and yield implications	**PROBLEM: Grass weeds in cereal crops like wheat are especially challenging to detect because they visually resemble the crop canopy, making traditional approaches unreliable. This paper investigates whether UAV-based imaging combined with deep learning could solve this problem.** The researchers deployed UAVs equipped with multispectral and hyperspectral sensors to collect aerial imagery. They applied semantic segmentation models such as U-Net and DeepLabv3 to classify weeds and crops at the pixel level. The final outputs were georeferenced weed maps that could be directly used for precision spraying.	The deep learning models achieved pixel-level accuracies above 85%, producing highly detailed weed maps that showed practical value for automated crop management. The study demonstrated that UAV-to-action pipelines, where drones detect weeds and trigger site-specific spraying, are feasible.	The models were computationally demanding, limiting their real-time application in the field. Dimensionality reduction and band selection were also necessary to handle the large hyperspectral datasets. Moreover, performance decreased in dense, mixed canopies, showing that scalability to more complex scenarios is still a challenge.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
6.	Oscar Leonardo Garcia-Navarrete, Li et al.	2024	Application of Convolutional Neural Networks in Weed Detection and Identification: A Systematic Review	**PROBLEM: While many CNN models have been proposed for weed detection, farmers and engineers lack clear guidelines on which models to use in practice, especially when hardware limitations are considered. This meta-analysis evaluates CNN-based weed detection studies to draw practical insights.** The authors systematically reviewed papers that applied CNNs for weed detection. They compared object detection models like YOLOv5 and Faster R-CNN with semantic segmentation models like U-Net, and also analyzed their computational demands for deployment on GPUs versus embedded devices.	The study found that CNNs consistently outperformed traditional machine learning methods, with YOLO-type detectors excelling in real-time detection scenarios and U-Net–based architectures delivering higher accuracy for pixel-level weed segmentation. However, speed and hardware requirements often dictated the choice of model in practice.	Despite these promising results, the meta-analysis noted that hyperspectral data was rarely integrated into CNN pipelines, leaving untapped potential. Another recurring issue was the small size of available datasets, which increases the risk of overfitting and reduces the generalizability of CNN-based solutions.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
7.	Garibaldi-Márquez et al.	2025	Advances on deep learning for proximal image-based weed recognition and control under authentic farmlands: A state-of-the-art review	This work is a **systematic literature review**. The authors analyzed 349 relevant papers published between 2015 and 2025 to summarize the state-of-the-art in deep learning for weed recognition. They categorized the findings by task (classification, detection, segmentation) and identified the most commonly used model architectures, noting that **ResNet/VGG** are popular for classification, **UNet** for segmentation, and the **YOLO family** for detection.	The review provides a comprehensive overview of the entire research field. Key contributions include: • A quantifiable analysis showing that the **YOLO family is the dominant architecture for weed detection**, appearing in **53.48%** of the reviewed detection studies. • A compilation of 34 publicly available crop-weed datasets, creating a valuable resource for researchers. • A specific focus on 26 studies that have deployed their models on real-world "smart weeders," bridging the gap between theory and practice.	The review identifies several critical gaps for the entire field of study: • **Lack of Robustness:** The primary challenge is that most models are trained on datasets that lack real-world variability (e.g., diverse lighting, growth stages, weed densities), which severely limits their ability to generalize and work reliably in different farm environments. • **Lab-to-Field Gap:** Many trained models are only evaluated on a test set of images and never deployed on actual smart equipment, where real-time performance and environmental factors become critical.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
8.	Mahrin Tasfe et al.	2025	Deep Learning-Based Models for Paddy Disease Classification and Segmentation: An Experimental Review	This paper is a **systematic experimental review** of deep learning models for paddy disease diagnosis. The authors did not just summarize papers; they tested **seven different classification models** (like DenseNet, ViT, MobileNet) and **eight different segmentation models** (like UNet, DeepLabv3+, TransUNet) on paddy disease datasets. Their method involved comparing the performance, size, and limitations of all these models to find the most suitable ones for real-world, resource-constrained applications.	The study provides a direct, hands-on comparison of numerous state-of-the-art models for agricultural disease analysis. Key contributions include: • Creating a **new, open-access dataset** for paddy disease segmentation to address a major resource gap. • Concluding that the **Deep Residual UNet** is the most suitable segmentation model, balancing performance and efficiency. • Demonstrating that traditional models like **DenseNet121** and ensemble models can achieve performance comparable to computationally expensive Vision Transformers (ViTs), making them better for on-farm applications.	The review identifies several critical gaps in the field of paddy disease diagnosis: • A significant lack of large, publicly available **segmentation datasets** for paddy diseases. • A lack of datasets for **paddy field weed segmentation** specifically. • Insufficient research into **disease severity analysis**, which requires precise segmentation. • A general need for models that are **memory-efficient** and can run on devices with limited computational power, like mobile phones for farmers.

S.NO	AUTHORS	YEAR	TITLE	METHOD	KEY CONTRIBUTIONS	GAP IDENTIFIED
9.	Armaano Ajay et al.	2025	An Explainable Deep Learning Framework for Sorghum Weed Classification Using Multi-Scale Feature Enhanced DenseNet	This paper focuses on **image-level classification**. The authors built a novel, custom deep learning architecture to classify images of sorghum and weeds. Their method uses a **DenseNet-169** model as a base and enhances it with two newly designed blocks (**DCFF** and **DMKCF**) for multi-scale feature extraction, plus a **Triplet Attention** module to help the model focus. They also used explainability techniques (Grad-CAM, LIME) to visualize the model's decision-making process.	The study makes several novel contributions to the field of weed classification: • It proposes a new, complex architecture that combines several advanced techniques, achieving a state-of-the-art **classification accuracy of 99.07%** on the SorghumWeedDataset. • It places a strong emphasis on **explainability (XAI)**, using tools like Grad-CAM to provide visual insights into why the model makes a certain prediction, which builds trust in the system.	The most significant gap is that this model is for **classification only**. It can accurately tell you *if* an image contains a "Broad-leaf weed," but it provides **absolutely no location information**. It cannot draw a bounding box or a segmentation mask to show *where* the weed is in the image. This makes it unsuitable for any real-world precision agriculture task that requires site-specific action, such as robotic weeding or targeted herbicide spraying.

RESEARCH CHALLENGES

- **High Visual Similarity:** A significant challenge is the strong visual resemblance between grassy weeds and the Sorghum crop itself, especially during early growth stages. They share similar colors, thin shapes, and textures, requiring a highly discriminative model to tell them apart.
- **Complex and Variable Field Environments:** Agricultural images captured by drones are not taken in a controlled lab. A robust model must overcome challenges like inconsistent lighting (strong sun and hard shadows), overlapping plants (occlusion), and different soil backgrounds, all of which can confuse a model.
- **Achieving Pixel-Level Precision:** Moving beyond simple object detection (which provides coarse bounding boxes) to instance segmentation (which provides precise outlines) is a major technical challenge. It requires a more sophisticated model and more detailed data to learn the exact boundaries of every plant, especially when they are clustered together.
- **Balancing Accuracy and Speed:** For any practical on-farm application, such as a drone or robot, a model must be both accurate and extremely fast. A key challenge is to incorporate advanced, performance-boosting features (like attention modules and complex augmentations) without sacrificing the real-time inference speed that models like YOLO are known for.

RESEARCH OBJECTIVES

- **To Develop a High-Precision Segmentation Model:** To implement and train a state-of-the-art deep learning model, YOLOv8-seg, for the task of instance segmentation, aiming to accurately identify and outline individual crops (Sorghum) and weeds at the pixel level.
- **To Investigate and Implement Novel Improvements:** To enhance the baseline model's performance by systematically introducing advanced techniques, including extensive data augmentation and architectural modifications like the integration of a CBAM attention module.
- **To Quantitatively Evaluate Model Performance:** To rigorously assess and compare the performance of the baseline and improved models using standard computer vision metrics, with a primary focus on **Mask mean Average Precision (mAP)**, to provide a clear measure of success.
- **To Demonstrate Superiority Over Existing Methods:** To highlight the practical advantages of the instance segmentation approach by comparing its precise, pixel-level output against the less detailed bounding-box-based detection methods found in the relevant literature.

PROPOSED ARCHITECTURE AND METHODOLOGY

Title: *Weed Detection and Segmentation using Deep Learning (YOLOv8) on RGB Agricultural Dataset*

Input: RGB images of crops + weeds (from Sorghum datasets).
Task: Train a model to automatically detect/segment weeds.
Output: Predicted segmentation masks showing where weeds are.

For ImageWeeds and Sorghum datasets, researchers have already used **YOLO (v3–v5)**, **Faster R-CNN**, **Mask R-CNN**, and **U-Net**.

PIPELINE USED:

1. Dataset Collection: Use ImageWeeds + Sorghum datasets (RGB images with annotations in JSON/XML/COCO/CSV).

2. Data Preprocessing: Convert all annotations into **YOLOv8 segmentation format** (polygon masks).

Perform **data augmentation** (flipping, rotations, brightness changes).

3. Model Selection: Choose **YOLOv8-seg** (segmentation version).

4. Training

Train YOLOv8-seg on your dataset.

Use train/val/test split (We have used 8:1:1)

5. Evaluation

Metrics: mAP (mean Average Precision), IoU (Intersection over Union), F1-score.

Compare results with YOLOv4 baseline paper.

6. Improvement

- **Model Architecture and Novel Enhancements (notimplemented yet)**

Attention Mechanism Integration: A **Convolutional BlockAttention Module (CBAM)** was integrated into the neck of theYOLOv8-seg network. This was accomplished by creating a custommodel configuration (.yaml) file that inserts the CBAM layer. Thepurpose of the CBAM is to enhance the model's featurerepresentation by allowing it to intelligently focus on the mostsignificant channels and spatial regions of an image, therebyimproving its ability to distinguish between the visually similarfeatures of crops and weeds.

- **Advanced Data Augmentation:** To improve the model's ability to generalize across varied, real-world field conditions, a strong set of data augmentations was applied on the fly during training. This included random rotations (degrees=20), scaling (scale=0.1), translation (translate=0.1), and color-space adjustments for saturation and value (hsv_s=0.5, hsv_v=0.5). Furthermore, the copy_paste augmentation was used to create more complex and challenging training scenes.
- **Learning Rate Scheduling:** A **Cosine Annealing Learning Rate Scheduler** (cos_lr=True) was employed. This method adjusts the learning rate following a cosine curve over the training epochs.
- **Class Imbalance Handling:** To address the underrepresentation of the "Grasses" class in the dataset, the classification loss weight was increased (cls=0.75). This technique forces the model to place a higher importance on correctly classifying all classes, especially the rarer ones, thereby improving overall model balance.
- Usings YOLOv8's built-in "Tuner" to automatically experiment with dozens of settings (like learning rate,momentum, etc.) to find the absolute best combination for specific dataset.

7. Result & Contribution

Show weed masks predicted by YOLOv8.

Highlight **performance improvement vs baseline paper**

PSEUDOCODES FOR IMPLEMENTATION

YOLOv8 Colab Pipeline:

- Install YOLOv8:

  pip install ultralytics
- Import YOLO:

  from ultralytics import YOLO
- Load pretrained YOLOv8 model (detection or segmentation):

  model = YOLO("yolov8n.pt") # nano model for speed
- Train on your dataset:

  model.train(data="sorghum.yaml", epochs=50, imgsz=640)

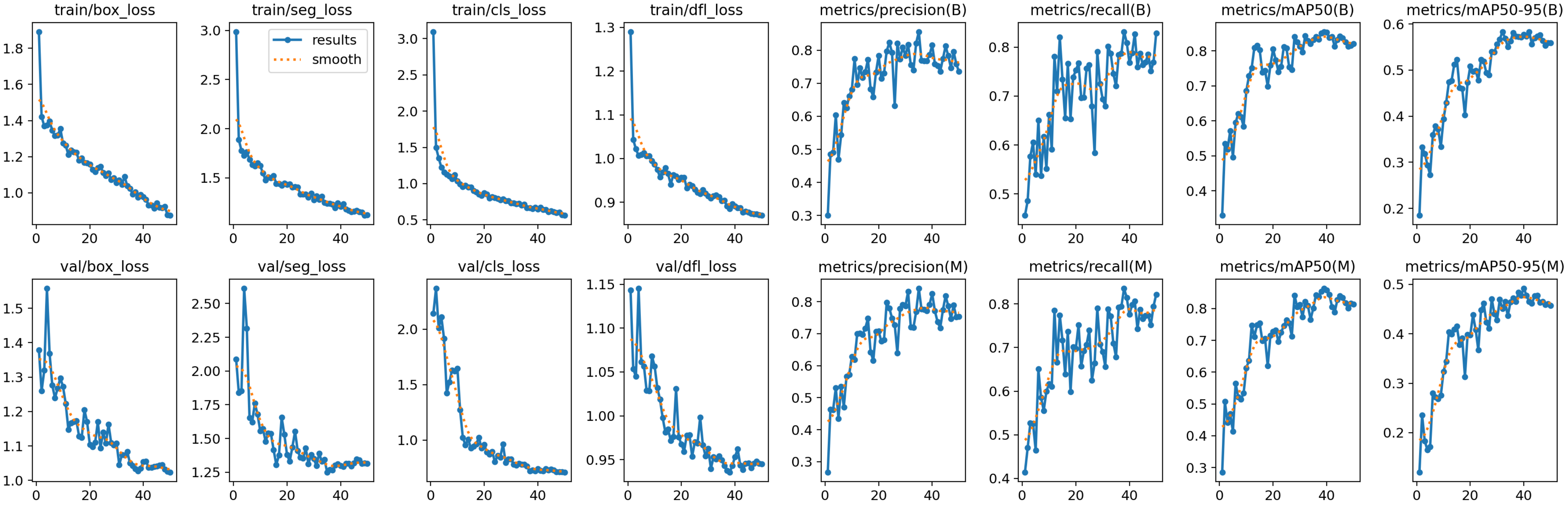
    - sorghum.yaml = config file that tells YOLO where your images + labels are.
- Evaluate results (mAP, precision, recall).
- Predict on test images:

  results = model.predict("path/to/image.jpg")

  results.show()

RESULTS,DISCUSSION AND COMPARISON

Key Result: The trained YOLOv8-seg model achieved an overall accuracy of **85.8% (Mask mAP@0.5)** on the validation dataset.



- **Insight 1: The Visual Similarity Challenge:** The slightly lower performance on "Grasses" is a key finding. This is due to the high visual similarity between thin, grassy weeds and the Sorghum crop itself, especially at certain growth stages. This is a common and difficult challenge in precision agriculture.
- **Insight 2: Precision and Stricter Metrics:** When evaluated at a stricter threshold (mAP@50-95), the overall mask accuracy was **49.3%**. This indicates that while the model is excellent at general identification, achieving pixel-perfect mask alignment for every instance is a significantly harder task.

```
Ultralytics 8.3.202 🚀 Python-3.12.11 torch-2.8.0+cu126 CUDA:0 (Tesla T4, 15095MiB)
YOLOv8s-seg summary (fused): 85 layers, 11,780,761 parameters, 0 gradients, 39.9 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95)     Mask(P          R      mAP50  mAP50-95): 100%
                   all         25        188      0.815      0.769      0.853      0.577      0.825      0.776      0.858      0.493
               Sorghum         24         91      0.881      0.734      0.875      0.625      0.894      0.744      0.884      0.583
               Grasses          9         16      0.738       0.75      0.786      0.474      0.742       0.75      0.786       0.37
      Broad-leaf weeds         11         81      0.826      0.822      0.898      0.634      0.838      0.833      0.904      0.525
Speed: 0.2ms preprocess, 5.6ms inference, 0.0ms loss, 5.7ms postprocess per image
Results saved to /content/drive/MyDrive/Sorghum_Project/training_runs/run_1_50_epochs
```

Grasses 0.76
Grasses 0.78
Broad−leaf weeds 0.86
Broad−leaf weeds 0.88
Grasses 0.82
Broad−leaf weeds 0.30
Broad−leaf weeds 0.44
Broad−leaf weeds 0.90
Sorghum 0.91
Grasses 0.65
Broad−leaf weeds 0.84
Broad−leaf weeds 0.73
Grasses 0.77
Sorghum 0.86
Broad−leaf weeds 0.86
Sorghum 0.61
Sorghum 0.46
Grasses 0.87
Sorghum 0.89
Broad−leaf weeds 0.44
Sorghum 0.92
Grasses 0.80
Sorghum 0.76
Broad−leaf weeds 0.88
Grasses 0.68
Broad−leaf weeds 0.79
Broad−leaf weeds 0.77
Broad−leaf weeds 0.50
Sorghum 0.25

- **Genze et al. (2022) - *Weed segmentation in sorghum fields***

Method: This paper performs **Semantic Segmentation** (classifies general pixel areas) on Sorghum fields using a **UNet/ResNet-34** model.

Performance: They reported a very high **F1-score of over 89%**.

Our Advantage: Our model achieved **85.8% mAP@50** on the more complex task of **Instance Segmentation** (outlining *individual* weeds). While the metrics aren't directly comparable, both models show high performance. Our model provides more granular, object-specific data, which is a significant advantage.

- **Zhao et al. (2022) - *Improved YOLOv4 in potato fields***

Method: This paper focuses on **Object Detection** (drawing **bounding boxes**) using a heavily optimized, lightweight **MC-YOLOv4**.

Performance: Their specialized model achieved an exceptional **98.52% mAP** for *detection*.

Our Advantage: Our model's **85.8% mAP** for *segmentation* is an excellent result that provides far more detail (precise masks vs. simple boxes). Furthermore, our model is significantly faster (**5.6 ms** vs. their **12.49 ms**).

- **Gallo et al. (2023) - *Performance of YOLOv7 on UAV images***

Method: This paper benchmarks **YOLOv7** for **Object Detection (bounding boxes)** on a new, challenging dataset.

Performance: They reported **56.6%** mAP@0.5 on their primary dataset.

Our Advantage: Our **YOLOv8** model achieved a much higher score of **85.8%** mAP@0.5 while performing the more complex task of **segmentation**.

REFERENCES

1. https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2020.611622/full?utm_
2. https://pmc.ncbi.nlm.nih.gov/articles/PMC8847227/?utm_
3. https://www.mdpi.com/2073-4395/14/2/363?utm_
4. https://www.sciencedirect.com/science/article/pii/S2666154324004253?utm_
5. https://www.mdpi.com/2072-4292/16/15/2808?utm_
6. https://plantmethods.biomedcentral.com/articles/10.1186/s13007-024-01272-6?utm_
7. https://www.mdpi.com/2077-0472/14/4/568?utm_
8. https://www.cambridge.org/core/journals/weed-science/article/abs/spectral-discrimination-of-crops-and-weeds-using-deep-learning-assisted-by-wavelet-transform-and-statistical-preprocessing/3DAC6F60EE437FF8C17740AA70E6F27A
9. https://www.mdpi.com/2079-9292/11/22/3709?utm_
10. https://www.mdpi.com/2072-4292/15/2/539?utm_
11. https://www.mdpi.com/2072-4292/16/23/4394?utm_
12. https://www.sciencedirect.com/science/article/abs/pii/S1161030125001480
13. https://www.sciencedirect.com/science/article/pii/S0168169922006962