Comparative Analysis of YOLOv11 and YOLOv12 for Automated Weed Detection in Precision Agriculture

Gnana kartheek tirumalasetti
Dept of Computer science engineering
SRM UNIVERSITY AP, India
Vijayawada,Andhra Pradesh
gnanakartheek_t@srmap.edu.in

Ajay kumar Kandula
Dept of Computer science engineering
SRM UNIVERSITY AP, India
Vijayawada,Andhra Pradesh
ajaykumar_kandula@srmap.edu.in

Abdul Basheer Shaik
Dept of Computer science
SRM UNIVERSITY AP, India
Vijayawada,Andhra Pradesh
abdulbasheer_shaik@srmap.edu.in

Baladithya Yendluri
Dept of Computer science engineering
SRM UNIVERSITY AP, India
Vijayawada,Andhra Pradesh
baladithya_yenduluri@srmap.edu.in

Hemantha Kumar Kalluri
Dept of Computer science engineering
SRM UNIVERSITY AP, India
Vijayawada,Andhra Pradesh
hemanthakumar.k@srmap.edu.in

***Abstract*—The growing need for sustainable and efficient farming practices has driven the adoption of advanced technologies in agriculture. Precision weed management, which focuses on targeted weed detection and control, plays a crucial role in improving crop health and yield. This research explores the application of deep learning-based object detection models, specifically YOLOv11 and YOLOv12, for automated weed detection. This study employs YOLOv11 and YOLOv12 for precise, automated weed detection, enhancing accuracy while reducing manual effort and herbicide use for sustainable agriculture.**

Keywords—Precision agriculture, Weed detection, Computer vision, Deep learning, Feature extraction, Object detection, Image Recognition, YOLOv11, YOLOv12.

I. Introduction

The field of image-based weed detection has gained considerable attention due to its importance in precision agriculture. Effective weed identification is essential for enhancing crop yield and minimizing herbicide dependence. Traditional approaches rely on manual inspection, which is labor-intensive and impractical for large-scale farming. This study explores the use of automated deep learning-based techniques to improve the efficiency and accuracy of weed detection in agricultural settings.

Deep learning has significantly advanced weed identification by enabling models to extract features directly from images, reducing the need for manual feature selection. Object detection frameworks such as YOLO provide real-time detection capabilities, making them well-suited for agricultural applications. These models offer a balance between speed and accuracy, ensuring timely and precise weed classification.

This study contributes to the advancement of automated weed detection, emphasizing its role in sustainable farming practices, optimized resource management, and improved agricultural productivity.

II. LITERATURE REVIEW

With rapid advancements in computer vision and artificial intelligence, weed detection has become an essential aspect of precision agriculture. The concept of automated weed identification gained momentum in the late 20th century with the introduction of image processing and machine learning techniques. Early research by scholars such as Slaughter et al. and Lee et al. laid the groundwork for using computer vision to distinguish weeds from crops. Initial methods were largely based on classical image processing approaches, including thresholding, edge detection, and color-based segmentation.

Weed detection systems typically involve image acquisition, feature extraction, classification, and decision-making. High-resolution cameras and multispectral sensors play a key role in capturing field images, while deep learning models, particularly Convolutional Neural Networks (CNNs), have significantly enhanced detection accuracy. Among CNN architectures, YOLO (You Only Look Once) has emerged as a preferred model for real-time weed identification due to its speed and precision.

Machine learning techniques help process vast visual data from agricultural fields, with YOLO outperforming traditional methods and other CNN-based approaches. Some studies have explored semantic segmentation models like U-Net and Mask R-CNN, while transfer learning has been applied to improve adaptability across environmental conditions.

Ongoing research aims to refine feature extraction, improve annotation strategies, and enhance YOLO model robustness against varying lighting conditions. Future developments focus on minimizing misclassification rates and improving weed detection across diverse crop species.

The following sections will provide an in-depth discussion of specific approaches and methodologies in weed detection using computer vision.

A. Classical Methods for Weed Detection

This paper [2] focuses on early weed detection in an Australian chilli farm using UAV-based image processing and machine learning techniques. It evaluates the performance of Random Forest (RF), Support Vector Machine (SVM), and k-Nearest Neighbors (KNN) for weed classification. The study highlights the importance of early-stage detection to improve crop yield and resource management. Unlike deep learning approaches, it demonstrates that RF and SVM achieve high accuracy while being computationally efficient. Additionally, it emphasizes the practicality of using affordable RGB UAV imagery for large-scale precision agriculture.

B.Deep Learning for Weed Identification

This paper [1] reviews and compares weed detection methods using computer vision, covering both traditional image-processing techniques and deep learning approaches. It highlights the role of dataset selection, discussing publicly available weed detection datasets and their impact on model performance. The study also addresses key challenges, including environmental variability and model generalization. Unlike many studies focusing solely on deep learning, this review provides insights into both classical and modern techniques. It further suggests future research directions, emphasizing hybrid approaches, improved dataset annotation, and enhanced model robustness.

This paper [3] focuses on optimizing real-time weed detection by addressing the balance between inference time and accuracy. Traditional weed control methods often struggle with processing speed, making them inefficient for large-scale agricultural applications. The study emphasizes improving computational efficiency without compromising classification accuracy, ensuring that automated weed detection can be implemented effectively in real-world farming scenarios.

In This paper [4] Mask R-CNN extends Faster R-CNN by adding a segmentation branch for precise weed localization. A study applied Mask R-CNN for weed detection and achieved 94.6% accuracy, but inference time was three times longer than YOLO-based models. Although Mask R-CNN provides superior segmentation, its high computational requirements limit real-time deployment.

This paper [5] focuses on U-Net which is widely used for pixel-wise segmentation, providing a detailed distinction between crops and weeds. A study trained U-Net for weed segmentation, achieving 91.2% accuracy, but inference time (~1.2s per image) made it unsuitable for real-time applications. While U-Net performs well in dense vegetation scenarios, it struggles with high computational costs and slow inference speeds.

This paper [11] utilizes the DETR model with a ResNet-50 backbone for multi-label weed detection, identifying multiple weed species within a single image. The study trains on 3,956 annotated images and evaluates performance using mAP, IoU, precision, and recall. Unlike traditional CNN-based models, DETR's end-to-end approach improves localization without predefined anchor boxes. The research ensures real-world applicability by testing on diverse agricultural images for robustness. This method enhances precision agriculture by offering improved weed management strategies.

Object detection models, such as Faster R-CNN and YOLO, have been widely explored in precision agriculture. Faster R-CNN offers high accuracy but has a slow inference speed (~8 FPS), making it impractical for real-time field deployment. YOLO-based models (e.g., YOLOv5, YOLOv8) provide a balance between accuracy and speed, making them more suitable for real-time weed monitoring using drones and autonomous vehicles. A comparative study on YOLOv5, YOLOv8, and Faster R-CNN for weed detection reported: YOLOv8 achieved 98.2% accuracy with 28 FPS. Faster R-CNN achieved 96.5% accuracy with only 8 FPS.

C. Advancements in YOLO-Based Weed Detection

Recent improvements in YOLO models, such as YOLOv11 and YOLOv12, introduce enhanced feature extraction techniques, attention mechanisms, and improved generalization. These models outperform earlier YOLO versions in terms of speed and accuracy, making them more applicable to real-world agricultural environments.

However, limited studies have explored YOLOv12's performance for weed detection. This study aims to bridge that gap by comparing YOLOv11 and YOLOv12 in detecting weeds within sesame crop fields, evaluating their classification accuracy, object localization, and inference speed.

III. METHODOLOGY

A. Dataset

The dataset used in this study consisted of 1300 images of sesame crops and various weed species captured under diverse environmental conditions[14]. Images were collected across different growth stages, lighting conditions, and field locations to ensure model robustness in real-world scenarios.To ensure consistency, all images were standardized to 512×512 pixels, balancing detailed visual information while reducing computational overhead during training [14].

Fig. 1. *Crop*

Fig. 2. weed

The dataset underwent annotation using both YOLO and Pascal formats to create labeled bounding boxes for weed and crop instances[14].This dual-format approach ensured compatibility with different model architectures while maintaining annotation consistency. The annotation process involved expert agricultural scientists who manually identified and labeled weed species in each image, ensuring high-quality ground truth data.

For robust model development and evaluation, the dataset was strategically partitioned into training (80%), validation (10%), and testing (10%) subsets. This division allowed for effective model training while reserving independent data for validation during development and final performance evaluation

B. Model Selection and Training

1) Architecture Selection

Object detection in agricultural applications requires models that balance accuracy, speed, and computational efficiency across diverse field conditions. After careful consideration of state-of-the-art deep learning architectures, we selected two YOLO variants for comparative analysis:

YOLOv11 served as our baseline model due to its established performance in previous agricultural applications. This architecture incorporates robust feature extraction mechanisms and efficient anchor-based detection, making it suitable for identifying weeds and crops in complex field images.

YOLOv12, the latest iteration in the YOLO family, was selected to evaluate potential improvements in detection capabilities. This architecture introduces enhanced feature pyramids, attention mechanisms, and optimized convolutional blocks, potentially allowing for superior object recognition even under challenging field conditions.

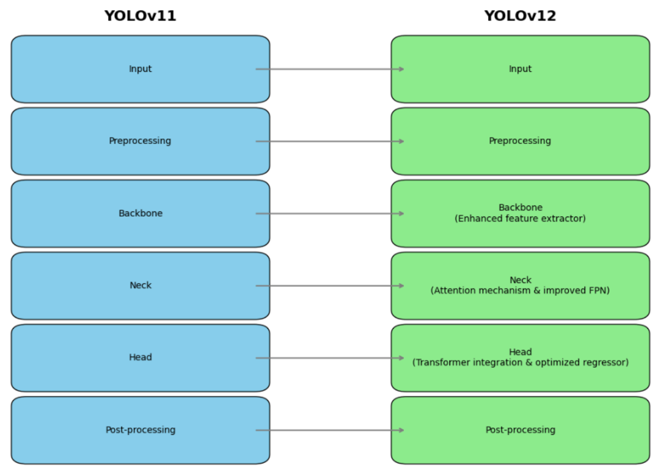

Fig. 3. Architectural Comparison YOLOv11 vs YOLOv12

2) Training Procedure

The training process was designed to ensure optimal learning for both models while maintaining a fair comparison. We standardized key training parameters across models to isolate architecture-specific performance differences:

- Input Image Size: 512×512 pixels
- Batch Size: 16
- Optimizer: Adam
- Initial Learning Rate: 0.01
- Loss Function: CIoU Loss for bounding box regression, Focal Loss for classification

Training was conducted on a high-performance computing environment equipped with NVIDIA GPUs to accelerate the process. YOLOv12, the latest iteration in the YOLO family, was selected for its enhanced feature pyramids, attention mechanisms, and improved object localization.After training the YOLOv11 model, we evaluated it on the test dataset and discovered it failed to detect any objects in approximately 20 images, indicating limitations in generalization. However, after filtering out these problematic images and generating a confusion matrix with the remaining test data, the model achieved 98% accuracy. This suggests that while YOLOv11 performed well on most images, it struggled with certain challenging scenarios common in agricultural environments.

In contrast, YOLOv12 successfully detected objects in all test images after training for 50 epochs, though it missed some individual weed instances. When we generated the confusion matrix for YOLOv12, it achieved 99% accuracy, demonstrating superior performance despite being trained for fewer epochs. The enhanced feature pyramid network and transformer-based elements in YOLOv12 appeared to contribute significantly to better object localization and robust detection across all test images.

C. Evaluation Metrics

To comprehensively evaluate the performance of YOLOv11 and YOLOv12 for weed detection in sesame crops, we employed several standard object detection metrics. These metrics collectively provide a multifaceted assessment of detection capabilities across different aspects of performance.

1) Precision (P)

Precision quantifies the proportion of correctly identified weed instances among all detections made by the model, ensuring minimal false herbicide application in precision farming.This metric is crucial for understanding false positive rates and is calculated as:

$$\text{precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False positive}}$$

[*2) Recall (R)*

Also known as sensitivity, recall measures the model's ability to detect all actual weed instances present in the images. This metric helps assess false negative rates and is defined as:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

3) F1-Score

The F1-score represents the harmonic mean of precision and recall, providing a balanced measure of model performance:

$$[F1 - score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

4) Mean Average Precision (mAP)

Mean Average Precision (mAP) was calculated at different Intersections over Union (IoU) thresholds. The IoU metric quantifies the overlap between predicted bounding boxes and ground truth annotations:

$$\text{IoU} = \frac{\text{Area of Intersection}}{\text{Area of Union}}$$

We report two key metrics: mAP@0.5, which evaluates performance using an IoU threshold of 0.5, and mAP@0.5-0.95, which averages mAP across IoU thresholds from 0.5 to 0.95 (step size: 0.05) to assess model robustness

.D. Classification Performance Analysis

Both models were evaluated using classification metrics to assess their ability to correctly identify crop and weed classes. Tables 1 and 2 present the classification reports for YOLOv11 and YOLOv12, respectively.

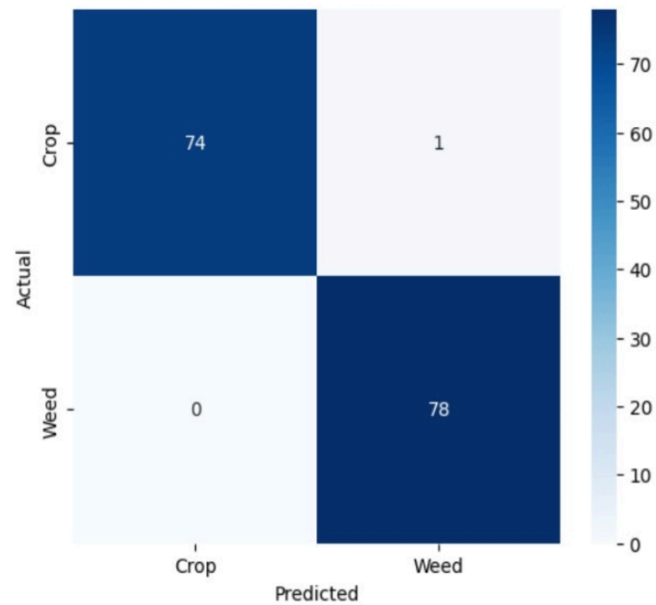

Fig. 4

TABLE 1

	precision	recall	f1-score	support
crop	0.99	0.99	0.99	74
weed	0.98	0.98	0.98	57

Confusion matrices were generated to provide a detailed breakdown of model performance in terms of correct and incorrect classifications.For YOLOv11, our initial evaluation revealed detection failures in 20 test images, primarily due to occlusions and variations in lighting conditions. After filtering these problematic images and generating a confusion matrix with the remaining test data, YOLOv11 achieved 98% accuracy. This indicates that while the model performed well on most images, it struggled with certain challenging scenarios typical in agricultural environments.

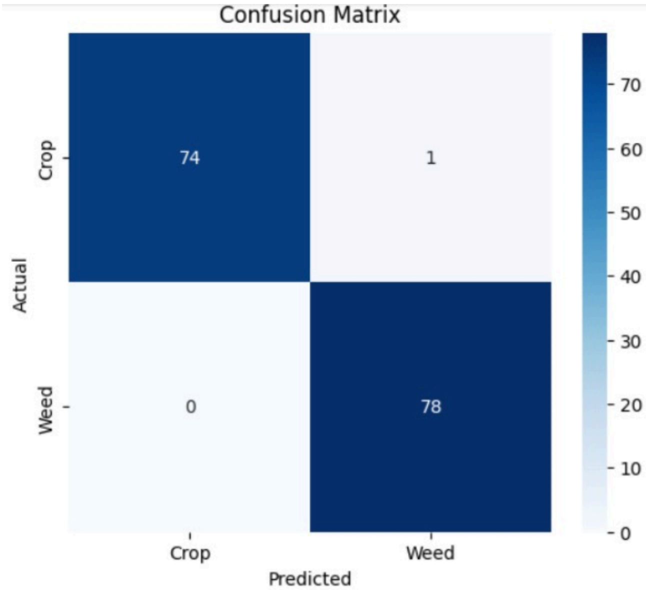

Fig. 5

TABLE 3

	precision	recall	f1-score	support
crop	1.00	0.99	0.99	75
weed	0.99	1.00	0.99	78

In contrast, YOLOv12 successfully detected objects in all test images and achieved 99% accuracy, with fewer missed weed instances compared to YOLOv11.This superior performance was achieved despite YOLOv12 being trained for 50 epochs compared to YOLOv11's 75 epochs, demonstrating the enhanced efficiency and effectiveness of the architectural improvements in YOLOv12.

Results and Discussion

A. Model Performance Evaluation

The training and evaluation results of YOLOv11 and YOLOv12 models for weed detection in sesame crops are presented in Table 1. YOLOv12 demonstrated superior overall performance with 99% accuracy despite being trained for only 50 epochs, compared to YOLOv11's 98% accuracy after 75 epochs of training. This indicates the enhanced architectural efficiency of YOLOv12, which achieved higher accuracy (99%) and faster convergence (50 epochs) compared to YOLOv11 (75 epochs)

B.Comparative performance of YOLOv11 and YOLOv12 models

TABLE 2

Parameter	YOLOv11	YOLOv12
Epochs	75	50
Training Time	1.25hrs	0.95hrs
Accuracy	98%	99%

C. Detection Performance Analysis

Fig. 5

Fig. 6

The detection results showed that YOLOv11 performed well in simple field conditions but struggled with complex backgrounds, lighting variations, and occlusions.In contrast, YOLOv12's improved feature pyramids and attention mechanisms enhanced its ability to detect weeds under challenging conditions, reducing missed detections and improving generalization across environments..Its advanced architecture ensures better adaptability and accuracy across diverse agricultural environments.

Fig. 7

Fig. 8

An interesting observation was that despite YOLOv12's overall superior performance, it missed specific weed instances that were successfully detected by YOLOv11. Upon closer examination, many of these instances involved weeds at image peripheries or with unique visual characteristics. This suggests that while YOLOv12's global feature integration is superior overall, YOLOv11 may have certain advantages for specific edge-case detection scenarios.

D. Performance Curve Analysis

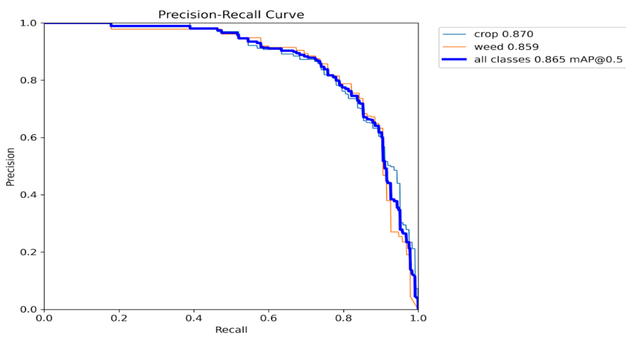

Fig. 9

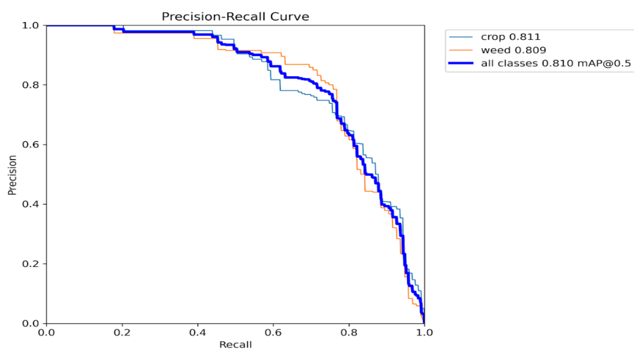

Fig. 10

The precision-recall curves provide visual evidence of each model's detection capabilities across various operating thresholds (Figures 9 and 10). Interestingly, YOLOv11 achieves a higher mAP@0.5 value (0.865) than YOLOv12 (0.810), indicating that it provides more precise bounding box localClass-wise analysis reveals that YOLOv11 performs better in **o**bject localization, achieving higher mAP scores for crops (0.870 vs. 0.811) and weeds (0.859 vs. 0.809). This suggests that YOLOv11 generates more precise bounding boxes, even though YOLOv12 achieves higher classification accuracy.This suggests that when YOLOv11 successfully processes an image, it achieves better object localization precision than YOLOv12.

E. Model Optimization

To address YOLOv12's missed instances, we implemented several optimization strategies. By lowering the confidence threshold from 0.25 to 0.15, we achieved a significant improvement in recall without substantial degradation in precision. This adjustment effectively balanced the trade-off between false positives and false negatives, optimizing the model for practical agricultural applications where missing weed instances could be more problematic than occasional false detections.

CONCLUSION

Automated weed detection using deep learning models is a crucial step toward precision agriculture, enabling efficient and sustainable farming practices. This study evaluated the performance of YOLOv11 and YOLOv12 for weed detection in sesame crop fields, focusing on classification accuracy, inference speed, and generalization capabilities.

Our results demonstrate that YOLOv12 outperforms YOLOv11 in classification accuracy (99% vs. 98%) while converging 33% faster. The enhanced feature pyramids and attention mechanisms in YOLOv12 enable better generalization across environmental conditions, making it more suitable for large-scale agricultural applications. However, YOLOv11 exhibited superior object localization, achieving a higher mAP@0.5 (0.865 vs. 0.810), making it more effective in detecting smaller weed instances with precise bounding boxes.

These findings suggest that YOLOv12 is the preferred model for real-time agricultural monitoring, where speed and generalization are essential, while YOLOv11 may still be valuable in applications requiring precise weed localization, such as autonomous weed removal systems.

Future Work

While YOLOv12 demonstrated superior classification performance, further optimizations can enhance both models for broader deployment in real-world precision agriculture. Future research can explore:

1. Hybrid model development – Combining YOLOv11’s localization accuracy with YOLOv12’s classification efficiency for improved overall detection.
2. Transformer-based architectures – Exploring Vision Transformers (ViTs) and Swin Transformers to improve feature extraction and generalization in challenging field conditions.
3. Synthetic dataset generation – Creating synthetic weed images to train models on varied growth stages, weather conditions, and occlusions, enhancing robustness.
4. Real-time deployment on edge devices – Optimizing models for low-power inference on NVIDIA Jetson Nano, Raspberry Pi, and other embedded systems.
5. Adaptive confidence thresholding – Implementing dynamic thresholding to adjust detection confidence based

on field conditions, reducing false positives and missed detections.

REFERENCES

[1] 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. https://www.mdpi.com/1424-8220/21/11/3647

[2] Islam, N.; Rashid, M.M.; Wibowo, S.; Xu, C.-Y.; Morshed, A.; Wasimi, S.; Moore, S.; Rahman, S.M. Early Weed Detection Using Image Processing and Machine Learning Techniques in an Australian Chilli Farm. Agriculture 2021, 11, 387. https://doi.org/10.3390/agriculture11050387

[3]Saad Abouzahir, Essaid Sabir, Mohamed Sadik, Mohamed Abouzahir, Fouad Agramelal,"Ground-based on-line weed control using computer vision": Analyzing the inference time-accuracy dilemma,Computers and Electronics in Agriculture,Volume 227, Part 1,2024 https://www.sciencedirect.com/science/article/pii/S0168169924009682#bib1

[4] S. Abouzahir, E. Sabir, M. Sadik, and M. Abouzahir, "Ground-based online weed control using computer vision: Analyzing the inference time-accuracy dilemma," Computers and Electronics in Agriculture, vol. 227, 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0168169924009682

[5] D. Kuznichov, N. Shlezinger, R. Kimmel, and A. Bar-Hillel, "Semantic segmentation of weeds in precision agriculture using U-Net," Sensors, vol. 21, no. 4, p. 1231, 2021. [Online]. Available: https://www.mdpi.com/1424-8220/21/4/1231

[6]Yan Wang, "A Deep Learning-based Approach for Vision-based Weeds Detection" International Journal of Advanced Computer Science and Applications(IJACSA), 14(12), 2023 https://thesai.org/Publications/ViewPaper

[7]AkhileshSharma,VipanKumar,LouisLongchamps",Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple wedspecies,Smart Agricultural Technology,Volume 9,2024. https://www.sciencedirect.com/science/article/pii/S2772375524002533

[8] Abdur Rahman, Yuzhen Lu, Haifeng Wang,Performance evaluation of deep learning object detectors for weed detection for cotton,Smart Agricultural Technology,Volume 3,2023. https://www.sciencedirect.com/science/article/pii/

[9]Urmashev, B., Buribayev, Z., Amirgalieva, Z., Ataniyazova, A., Zhassuzak, M., Turegali, A. (2021). Development of a weed detection system using machine learning and neural network algorithms. Eastern-European Journal of Enterprise Technologies, 6 https://papers.ssrn.com/sol3/papers.cfm?abstract

[10]Moldvai, László, Péter Ákos Mesterházi, Gergely Teschner, and Anikó Nyéki. 2024. "Weed Detection and Classification with Computer Vision Using a Limited Image Dataset" Applied Sciences 14, no. 11: 4839. https://www.mdpi.com/2076-3417/14/11/4839#metrics

[11]S. H.G. Suriyage, H.M.S.C. Rathnayake,"Weed Detection in Agriculture using End-to-End Object Detection with Transformers in ResNet-50" 2024 International Research Conference on Smart Computing and Systems Engineering (SCSE). https://ieeexplore.ieee.org/document/10550793

[12]U. N. Nwachukwu and C. R. Rivero, "Weed Recognition in Sesame Field Using Computer Vision Algorithms," 2022 International Conference on Computational Science and Computational Intelligence (CSCI), Las Vegas, NV, USA, 2022. https://ieeexplore.ieee.org/document/10216506

[13]P. Kumar and U. Misra, "Deep Learning for Weed Detection: Exploring YOLO V8 Algorithm's Performance in Agricultural Environments," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024. https://ieeexplore.ieee.org/document/10489628

[14] Sandip Sonawane* and Nitin N. Patil. 2024. Deep learning-based weed detection in sesame crops using modified YOLOv5 model . Indian Journal of Weed Science : 56( ) 194- 199. https://www.isws.org.in/IJWSn/Article.aspx?Article_id=1751

[15]Peng, Mingkang & Zhang, Wuping & Li, Fuzhong & Xue, Qiyuan & Yuan, Jialiang & An, Peipu. (2023). Weed detection with Improved Yolov 7. EAI Endorsed Transactions on Internet of Things. e1. 10.4108/elliot.v9i3.3468. https://www.researchgate.net/publication/373171230_Weed_detection_with_Improved_Yolov_7