When YOLO Meets SAM: Data-Efficient Weed Density Estimation

Soniya[1,2], Nived Ambadipudi[3,2], Apurva Narayan[4],

[1]SciDentAi, Innovation Centre, Indian Institute of Science, Bengaluru, Karnataka – 560012, India
[2]International Centre for Applied Systems Science for Sustainable Development, 34 Eagle Hill, Cambridge, N3C 2C9, Ontario, Canada
[3]Purdue University, 610 Purdue Mall, West Lafayette, Indiana 47907, United States
[4]Department of Computer Science, The University of Western Ontario, 1151 Richmond St, London, N6A 3K7, Ontario, Canada

Abstract

Weed infestation is a persistent problem in agriculture, particularly in organic farming, where chemical herbicides are restricted and manual weeding is labour-intensive and costly. Accurate and automated weed identification is therefore critical for sustainable crop management, yet conventional detection and segmentation models demand extensive annotated datasets, making large-scale deployment impractical. To address this, the present study introduces a zero-shot instance segmentation framework that integrates the YOLO segmentation model with the Segment Anything Model (SAM). YOLO generates precise bounding boxes for weed regions, which SAM then refines to produce high-quality masks for accurate weed density estimation. This cooperative approach combines YOLO's strong localisation with SAM's fine-grained segmentation, achieving robust results even with minimal annotated data. The proposed framework significantly reduces annotation and computational costs while maintaining high accuracy, offering an efficient and scalable solution for precision weed management in agriculture.

Introduction

Weed identification in agriculture has become increasingly vital, especially given the global challenge of feeding a projected population of 10 billion by 2050 (United Nations Environment Programme 2024). Farmers face a significant challenge from the widespread proliferation of weeds, which causes substantial annual losses. Weeds compete with crops for essential nutrients, water, and sunlight, significantly reducing yields. Effective weed management strategies traditionally relied on understanding weed biology to cultivate crops with minimal weed interference. However, due to the uncertainty of weed locations, the indiscriminate application of herbicides across entire fields poses risks to crop yields and the environment.

By identifying specifically weed locations, farmers can implement targeted management strategies that reduce the use of broad-spectrum herbicides. This approach lowers costs and minimises environmental harm, promoting sustainable agricultural practices. Organic farming provides eco-friendly, sustainable methods, though it requires considerable effort to maintain crop health (Krause et al. 2024; Eyhorn et al. 2019). Controlling weeds remains one of the significant challenges in organic agriculture. Automation offers a solution that can significantly enhance sustainable organic farming practices to a great extent.

Weeds pose a significant threat to organic farming, and their timely identification is crucial for enhancing yield production. Manually identifying weeds becomes a tedious task when fields are large. Automation can significantly assist in weed identification through machine learning models. These models can quickly identify weed species and their areas in a field, allowing for accurate estimates of weed density. Farmers can focus solely on effective remedies to control weeds with this information.

Over the past decade, deep learning models have achieved remarkable success in various fields, including computer vision, healthcare, cybersecurity, and others (Ahmed et al. 2023; Sarker 2021). Researchers actively demonstrate the capabilities of machine learning in agriculture by tackling tasks such as disease, weed, and crop-related classification, detection, and segmentation problems (Attri, Awasthi, and Sharma 2024; Aashu et al. 2024). The similarity between crops and weeds makes detection and segmentation particularly challenging. Previous studies (Gallo et al. 2023; Narayana and Ramana 2023; Rehman et al. 2024; Liu et al. 2024; Guzel et al. 2024; Niu et al. 2024; Goyal, Nath, and Niranjan 2025) explored the feasibility of real-time weed detection using object detection algorithms. Meanwhile, other approaches (Genze et al. 2022; Charania et al. 2023; Kong et al. 2024b,a) utilised segmentation algorithms to pinpoint areas occupied by weeds. These studies used the You Only Look Once (YOLO) model as the underlying framework and thousands of annotated images for training. The quality and quantity of annotated datasets are key factors driving the success of YOLO models.

Data annotation presents an overarching challenge in these methodologies due to its time complexity and labour-intensive nature. To address this challenge, this study presents a data-driven fusion of the conventional segmentation process with a zero-shot instance segmentation method to reduce this labour-intensive process. This method efficiently calculates weed density even when a model is trained with a few annotated images. The methodology can identify weeds in any crop. This study effectively identified weeds in the pigeon peas field, demonstrating its practicality and efficiency.

The rest of the paper is organised as follows: Section de-

Figure 1: A few samples of the frames extracted from the collected videos

tails the procedure used to collect data. Section presents the methodology used in this study, and the corresponding experimental results are detailed in Section . A comparison of the proposed methodology with other models is presented in section . Section discusses various aspects of the methods used, and section serves as the paper's conclusion.

Data Collection

In this study, researchers collected data through mobiles from a farm located at coordinates 27°14'19.2"N 78°00'39.1"E, Agra, Uttar Pradesh, India. They employed the iPhone 12 mini and iPhone SE to gather video data from a height of approximately 35-37 inches from the ground. Next, they used the Roboflow platform to extract frames at one frame per second and meticulously annotate them manually. Additionally, they introduced noise, saturation, and blur into these images through augmentation techniques to enhance the robustness of our approach. Figure 1 shows a few sample images from the collected dataset.

The annotated frames are then used to examine the feasibility of existing state-of-the-art model namely SAM, whose details are discussed in the next setcion .

Methodology

This study aims to address the weed segmentation and associated weed density problem using state-of-the-art (SOTA) models such as the Segmentation Anything Model (SAM) and the You Only Look Once (YOLO) models. Before delving into the detailed exploration of various approaches integrating SOTA models, we briefly present preliminary results obtained using the SAM model for the segmentation task.

SAM Based Weed Segmentation

The Segmentation Anything Model (SAM) (Kirillov et al. 2023) is a cutting-edge segmentation model renowned for its zero-shot learning capability, which enables it to segment any object within an image. However, it did not perform well when we tested SAM with our dataset. As shown in figure 2, SAM segmented some objects in the image without clearly discriminating between weeds and pigeon pea plants. Since we focus specifically on segmenting weeds, SAM only needs to process the details of weed regions. Therefore, we integrated a detection model with SAM, which calculates and processes the bounding box coordinates to delineate the weed regions. SAM then uses these bounding boxes to segment the weeds contained within them accurately.

We need a model to accurately detect and highlight weed regions to estimate weed density in a field rather than processing the entire image as input to SAM. To achieve this,

(a) (b)

Figure 2: SAM's output on a sample image from our dataset shows that in (a), it segments some of the objects present in the image (including weeds and plants), whereas in (b), it segments the whole images based on their features. However, to isolate and segment only the area of interest, i.e., weeds, it's necessary first to detect the weeds separately and then apply SAM for segmentation specifically on those regions.

we considered using the You Only Look Once (YOLO) segmentation model (Solawetz and Francesco 2024), renowned for its real-time object detection and segmentation capabilities. The model's performance largely depends on the size of the annotated datasets used for training, meaning that reliable weed density evaluation is contingent on the availability of such datasets. However, creating large annotated datasets in real-time scenarios is a challenging task. To address this issue, we propose a hybrid SAM-based YOLO segmentation model that effectively performs with small and large datasets. The details of this hybrid model are presented in the next section .

Proposed Approach

This study presents a hybrid weed segmentation framework that effectively reduces the need to annotate large numbers of images while still accurately estimating weed density. The framework's basis is to detect the weed regions, generate bounding boxes around them, and segment the actual weed area within those boxes. State-of-the-art deep learning models, such as R-CNN (Girshick et al. 2013), Mask R-CNN(He et al. 2017), EfficientDet (Tan and Le), YOLO (Redmon et al. 2016), SAM (Kirillov et al. 2023), OMG-seg (Li et al. 2024), Florence (Yuan et al. 2021), Detectron (Ross Girshick and Ilija Radosavovic and Georgia Gkioxari and Piotr Dollár and Kaiming He 2018), and others, can be used to identify and segment weeds within this framework.

We considered the You Look Only Once (YOLO) segmentation model, which consists of a detection and segmentation module. The proposed framework optimises the baseline approach by carefully selecting the models for detection and segmentation tasks. We consider the YOLO segmentation model the underlying model of our study because it offers detection and segmentation, eliminating the need for two different models. Additionally, this model's lightweight architecture makes it suitable for deployment purposes.

The proposed weed segmentation framework is shown in Figure 3. It employs a YOLO segmentation model that includes both a detection and a segmentation module as its primary model. The model identifies weed-occupied regions in the field images through the detection module, highlights them with bounding boxes, and outputs their coordinates. This output, D_O^{YOLO}, serves as the input for the segmenta-

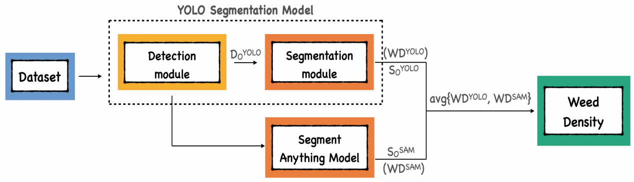


Figure 3: A pictorial representation of the proposed framework, which integrates the SAM model with the segmentation module of a YOLO-based segmentation model, for weed segmentation that works for small as well as large size datasets and helps in calculating weed density

tion module, which segments the weed area from the identified regions. Since YOLO-based detection and segmentation models perform well with large datasets, their performance declines significantly with smaller datasets (as discussed in section). To address this, we integrate the Segment Anything model to enhance the performance for smaller datasets, as illustrated in Figure 3. Now, the detection model's output D_O^{YOLO} will pass through both the segmentation module and SAM.

Since the models exhibit distinct characteristics, we cooperatively combine them to calculate weed density more effectively. As shown in Figure 3, we compute the weed density using the terms WD^{YOLO} and WD^{SAM}, which represent the weed density calculated by the standalone YOLO segmentation model and its integration with SAM, respectively. To evaluate the weed density predicted by a model, we sum the areas of the segmented regions and subtract any overlapping areas. The overall weed density is then assessed using the following equation 1

$$WD_{Avg} = \frac{\alpha WD^{YOLO} + (1-\alpha)WD^{SAM}}{2} \tag{1}$$

where, the term WD_{Avg} defines the average weed density. The parameter α represents a user-defined value that helps calculate a weighted average of weed density evaluated by the individual models. Since the YOLO segmentation model generates accurate masks, a high value of α yields a more precise weed density value. On the other hand, a lower value of α results in a weed density estimate based on masks generated by SAM, which are finer than those generated by the YOLO segmentation model.

Experimental observations show that integrating SAM with the YOLO model's segmentation module enhances overall performance, particularly in scenarios with limited annotated images. Details of these experiments are discussed in the next section .

Experiments and Results

We evaluate the models' performance in this approach in terms of mean average precision (mAP)[1]. To examine the effect of training set size on the YOLO segmentation model, we train it with varied training sets of 261, 200, 150, 100, and 50 images. The validation set comprises 100 images. Next, we evaluate the performance of the models in conjunction with SAM.

Table 1 presents the results of the YOLO segmentation model trained with different training set sizes. The table clearly shows that the model performs better when we train it with a significant number of annotated images and decreases as we reduce the training set size. This indicates that training the model is quite expensive because it requires a substantial number of annotated images, which poses a significant challenge regarding the time and effort needed for annotation. Next, we integrate SAM with the YOLO segmentation model to address this challenge and support the weed detection task.

Table 1: Effect of the training set's size on the approach based on the standalone YOLO segmentation model and its conjunction with SAM: performance is evaluated using the metric mAP@50

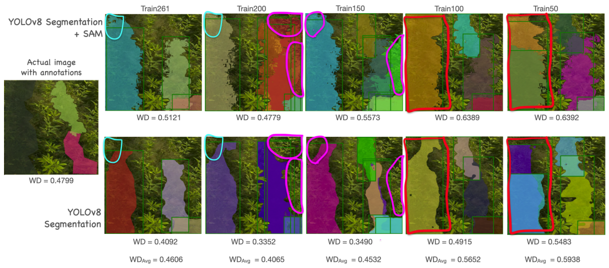
Images	YOLOv8 Segmentation (S_O^{YOLO})	YOLOv8 Segmentation + SAM (S_O^{SAM})
261	0.431	0.443
200	0.385	0.636
150	0.359	0.620
100	0.354	0.683
50	0.286	0.707

When we use the YOLO segmentation model in conjunction with SAM, we notice (as shown in table 1) a reverse effect compared to what we observed in the previous case (as discussed above). With limited training images, the YOLO model with SAM demonstrates superior performance compared to its standalone performance. Since the model is not finely trained with a limited number of images, it generates loose bounding boxes. In this case, SAM has a higher flexibility in segmenting weeds from plants within these boxes. On the other hand, as we increase the training set size, the model's performance improves, which generates accurate bounding boxes, limiting SAM's flexibility and decreasing overall performance.

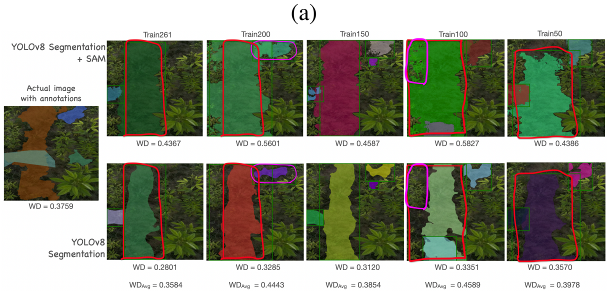
Figure 4 compares the inferences generated by the detection module of the YOLO segmentation model with those from its segmentation module and SAM. The right-most subfigures in figure 4a and 4b show that when the detection module trains on fewer annotated images, SAM performs better at segmenting weeds within the bounding boxes. It covers more area (marked in red) than the segmentation module. We observe the same fact in each subfigure of figure 4a and 4b.

On the other hand, when the detection module is trained on a significant number of annotated images, the YOLO model's segmentation module outperforms SAM by more accurately capturing the weed-occupied regions. It has been observed that SAM sometimes cannot differentiate between weeds and plants, mistakenly categorising plants as weeds. This is highlighted in the magenta markings in the 2nd and 3rd subfigures from the left in Figure 4a and in the 2nd and

[1]The mAP score indicates overall accuracy, calculated by averaging the precision versus recall graph by computing the area under the curve



(a)



(b)

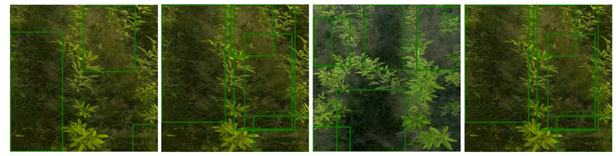
Figure 4: Comparing inferences generated by the detection module of the YOLO segmentation model followed by its segmentation module and SAM: (a) SAM generates finer masks, while (b) the YOLO model's segmentation module provides more accurate masks.

4th subfigures from the left in Figure 4b. Additionally, the cyan markings in the first two leftmost subfigures of figure 4a indicate the areas where SAM incorrectly identifies plants as weeds. In contrast, the YOLO model accurately detects and excludes the plant from the segmentation mask, even when the manual annotator wrongly labels that area as a weed. This highlights the superior accuracy of the YOLO segmentation model in producing precise masks.

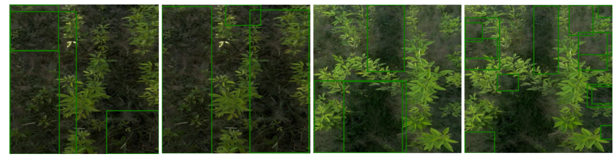
From Figure 4, it is clear that SAM segments weeds from plants with high precision in all cases, closely following the boundary between them. In contrast, the YOLO segmentation model maintains a safe margin from the plants, marking the weed-occupied region within the bounding box. This difference results in SAM generating finer, more detailed segmented areas.

This analysis shows that while SAM produces high-quality masks, the YOLO model's segmentation module delivers more accurate masks. Our proposed approach combines these strengths, generating precise and finely detailed masks that highlight weed regions better. This fusion ultimately contributes to a more reliable evaluation of weed density.

Our experiments showed that the two approaches—using the standalone YOLO segmentation model and combining the YOLO model with SAM—exhibit distinct characteristics. Table 1 shows that the standalone YOLO segmentation performs well with significant-size training datasets, but when combined with SAM, it performs better with smaller datasets. Moreover, the analysis in figure 4 reveals that each approach generates masks with different qualities: one excels at capturing fine details, while the other produces more accurate, high-quality masks. Given these differences, neither approach alone is sufficient for reliably calculating weed density. To address this, we propose a cooperative fusion of both techniques (as illustrated in figure 3), enabling the method to work effectively with small and large datasets.



(a)



(b)

Figure 5: Bounding boxes generated by detection models, (a) YOLOv8 and (b) YOLOWorld that often cover weeds and plants. As a result, these boxes cannot provide reliable weed density estimates.

Figure 4 illustrates the weed density captured by each approach and their averages. The SAM-based approach tends to generate precise masks and covers a larger area, resulting in a higher weed density value than the standalone YOLO segmentation model. However, because the SAM model occasionally fails to distinguish between plants and weeds, it mistakenly segments plants as weeds, artificially inflating weed density. On the other hand, the YOLO segmentation model calculates weed density more accurately. While the YOLO model generates masks with a safe margin around plants, it may not capture the actual weed density in certain areas. Therefore, as shown in Figure 3, we fuse the outputs of both models and use the average operator to estimate the total weed density more reliably.

As shown in Figures 4a and 4b, the calculated weed density is occasionally higher than the manually marked weed density in the actual images. This indicates that the proposed approach provides a reliable estimation of weed density in a given region, capturing areas that might otherwise be overlooked in manual annotations.

Comparison

Table 2 compares our proposed approach with other methods for estimating weed density values. The first column of the table displays manually annotated actual images, while the next column lists the corresponding weed density values. Equation 1 defines how our approach estimates the weed density value. We compare our approach with two other methods: i) standalone detection models and ii) their integration with SAM. This study uses the YOLOv8 and YOLOWorld detection models. Since these detection models generate bounding boxes around the most probable weed-occupied regions and sometimes, the generated boxes also enclose plants (as shown in figure 5), they cannot accurately estimate weed density. As a result, we mark $\times$ in table 2.

Table 2: Comparison of estimated weed density (WD) from randomly selected test images using models trained with 100 images. The proposed YOLO+SAM approach yields superior density estimation with $\alpha = 0.5$ in Eq. 1.

Input Image	Actual Value	Proposed Method	YOLOv8 + SAM	YOLOWorld + SAM
	0.4799	**0.5652**	0.5070	0.4288
	0.6502	**0.6531**	0.5226	0.4765
	0.3414	**0.4166**	0.2599	0.3939
	0.6782	**0.6581**	0.5570	0.6433
	0.3759	**0.4589**	0.1233	0.4104

We integrate SAM with detection models, specifically YOLOv8 and YOLOWorld, to estimate weed density. Table 2 compares this approach with our proposed approach. The table clearly shows that our proposed approach outperforms the other methods. The results are more consistent since our approach calculates the average weed density estimated by multi-models (as shown in figure 3). Additionally, our approach uses a segmentation model as its core, providing a reliable estimate of weed density. The SAM model further enhances performance by capturing both broader and finer areas. As a result, for most images, as shown in table 2, our calculated weed density is higher than the manually marked weed density in the actual images.

Discussion

Experimental results show that the proposed approach effectively handles weed segmentation tasks even with a few annotated images. This effectiveness stems from integrating SAM with the YOLO segmentation model, eliminating the need for an ample annotated training set to train the YOLO model. Additionally, SAM's zero-shot segmentation capability enhances the practicality of this integration.

In this study, we considered the YOLO segmentation model as our underlying model. This choice is due to the inclusion of both detection and segmentation modules, eliminating the need for a separate segmentation model when using the YOLO or YOLO-worldv2 detection models.

Instead of using a YOLO-based model, we can consider foundational models like Florence2 and Detectron. These models, trained on vast datasets, tackle more complex tasks than just detection or segmentation. Although they offer zero-shot learning, they did not perform well when we evaluated them with our dataset[2]. One of the possible reasons for this is the lack of domain knowledge. This finding highlights the need for fine-tuning these models. Our experiments show that fine-tuning requires a significant number of annotated images. As we used in this study, small datasets did not effectively train these models, indicating that fine-tuning foundational models necessitates comparatively large datasets.

Conclusion

We compare the work conducted in this study with similar research discussed in (Genze et al. 2022; Wang et al. 2022; Charania et al. 2023; Kong et al. 2024b,a; García-Navarrete et al. 2024; Liu et al. 2024; Guzel et al. 2024; C et al. 2024). These studies employed different models, such as YOLOv5, YOLOv8, and CNN, to segment or detect weeds and crops. They used thousands of annotated images to train their models, which enabled them to achieve high-performance metrics.

In contrast, our approach involved fine-tuning the YOLOv8 segmentation model's detection module for weed detection from pigeon pea fields and employing its segmentation module and SAM for weed segmentation. The integration of SAM helps us achieve high-performance metrics even in the case of a limited number of annotated images. Additionally, SAM's zero-shot knowledge transfer obviated the need for further fine-tuning. Choosing the YOLOv8 segmentation model eliminates the need for a separate segmentation model when using the YOLO or YOLO-worldv2 detection models. Thus cutting extra computation costs. Therefore, the proposed framework saved annotation efforts and reduced computational requirements.

Acknowledgements

The authors thank the International Centre for Applied Systems Science for Sustainable Development (ICASSSD) for arranging a field for data collection, as mentioned in this study. They also thank Arsh Chaudhary, a PhD candidate at Western University in Canada.

References

Aashu; Rajwar, K.; Pant, M.; and Deep, K. 2024. Application of Machine Learning in Agriculture: Recent Trends and Future Research Avenues. *ArXiv*, arXiv:2405.17465v.

Ahmed, S. F.; Alam, M. S. B.; Hassan, M.; Rozbu, M. R.; Ishtiak, T.; Rafa, N.; Mofijur, M.; Ali, A. B. M. S.; and Gandomi, A. H. 2023. Deep learning modelling techniques: cur-

[2] We trained the Florence2 model for 50 epochs with r=32 and 16. In both cases, its performance (mAP@50) was 0.28 for the detection task.

rent progress, applications, advantages, and challenges. *Artificial Intelligence Review*, 56: 13521–13617.

Attri, I.; Awasthi, L. K.; and Sharma, T. P. 2024. Machine learning in agriculture: a review of crop management applications. *Multimedia Tools and Applications*, 83: 12875–12915.

C, S. G.; Upadhyay, A.; Zhang, Y.; Howatt, K.; Peters, T.; Ostlie, M.; Aderholdt, W.; and Sun, X. 2024. Field-based multispecies weed and crop detection using ground robots and advanced YOLO models: A data and model-centric approach. *Smart Agricultural Technology*, 9: 100538.

Charania, S.; Lendave, P.; Borwankar, J.; and Kadge, S. 2023. A Novel Approach to Weed Detection Using Segmentation and Image Processing Techniques. In *World Conference on Communication & Computing (WCONF)*. IEEE.

Eyhorn, F.; Muller, A.; Reganold, J. P.; Frison, E.; Herren, H. R.; Luttikholt, L.; Mueller, A.; Sanders, J.; Scialabba, N. E.-H.; Seufert, V.; and Smith, P. 2019. Sustainability in global agriculture driven by organic farming. *Nature Sustainability*, 2: 253–255.

Gallo, I.; Rehman, A. U.; Dehkordi, R. H.; Landro, N.; Grassa, R. L.; and Boschetti, M. 2023. Deep Object Detection of Crop Weeds: Performance of YOLOv7 on a Real Case Dataset from UAV Images. *Remote Sensing*, 15(2).

García-Navarrete, O. L.; Santamaria, O.; Martín-Ramos, P.; Valenzuela-Mahecha, M. Á.; and Navas-Gracia, L. M. 2024. Development of a Detection System for Types of Weeds in Maize (Zea mays L.) under Greenhouse Conditions Using the YOLOv5 v7.0 Model. *Agriculture*, 14(2).

Genze, N.; Ajekwe, R.; Güreli, Z.; Haselbeck, F.; Grieb, M.; and Grimm, D. G. 2022. Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields. *Computers and Electronics in Agriculture*, 202: 107388.

Girshick, R. B.; Donahue, J.; Darrell, T.; and Malik, J. 2013. Rich feature hierarchies for accurate object detection and semantic segmentation. *CoRR*, abs/1311.2524.

Goyal, R.; Nath, A.; and Niranjan, U. 2025. Weed detection using deep learning in complex and highly occluded potato field environment. *Crop Protection*, 187: 106948.

Guzel, M.; Turan, B.; Kadioglu, I.; Basturk, A.; Sin, B.; and Sadeghpour, A. 2024. Deep learning for image-based detection of weeds from emergence to maturity in wheat fields. *Smart Agricultural Technology*, 9: 100552.

He, K.; Gkioxari, G.; Dollár, P.; and Girshick, R. 2017. Mask R-CNN. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 2980–2988.

Kirillov, A.; Mintun, E.; Ravi, N.; Mao, H.; Rolland, C.; Gustafson, L.; Xiao, T.; Whitehead, S.; Berg, A. C.; Lo, W.-Y.; Dollar, P.; and Girshick, R. 2023. Segment Anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2023, pp. 4015-4026*.

Kong, X.; Li, A.; Liu, T.; Han, K.; Jin, X.; Chen, X.; and Yu, J. 2024a. Lightweight cabbage segmentation network and improved weed detection method. *Computers and Electronics in Agriculture*, 226: 109403.

Kong, X.; Liu, T.; Chen, X.; Jin, X.; Li, A.; and Yu, J. 2024b. Efficient crop segmentation net and novel weed detection method. *European Journal of Agronomy*, 161: 127367.

Krause, H.-M.; Mäder, P.; Fliessbach, A.; Jarosch, K. A.; Oberson, A.; and Mayer, J. 2024. Organic cropping systems balance environmental impacts and agricultural production. *Scientific Reports*, 14(25537).

Li, X.; Yuan, H.; Li, W.; Ding, H.; Wu, S.; Zhang, W.; Li, Y.; Chen, K.; and Loy, C. C. 2024. OMG-Seg: Is One Model Good Enough For All Segmentation? *arXiv preprint,*, arXiv:2401.10229.

Liu, Y.; Zeng, F.; Diao, H.; Zhu, J.; Ji, D.; Liao, X.; and Zhao, Z. 2024. YOLOv8 Model for Weed Detection in Wheat Fields Based on a Visual Converter and Multi-Scale Feature Fusion. *Sensors*, 24(13).

Narayana, C. L.; and Ramana, K. V. 2023. An Efficient Real-Time Weed Detection Technique using YOLOv7. *International Journal of Advanced Computer Science and Applications(IJACSA)*, 14(2).

Niu, W.; Lei, X.; Li, H.; Wu, H.; Hu, F.; Wen, X.; Zheng, D.; and Song, H. 2024. YOLOv8-ECFS: A lightweight model for weed species detection in soybean fields. *Crop Protection*, 184: 106847.

Redmon, J.; Divvala, S.; Girshick, R.; and Farhadi, A. 2016. You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788.

Rehman, M. U.; Eesaar, H.; Abbas, Z.; Seneviratne, L.; Hussain, I.; and Chong, K. T. 2024. Advanced drone-based weed detection using feature-enriched deep learning approach. *Knowledge-Based Systems*, 112655.

Ross Girshick and Ilija Radosavovic and Georgia Gkioxari and Piotr Dollár and Kaiming He. 2018. Detectron. https://github.com/facebookresearch/detectron.

Sarker, I. H. 2021. Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. *SN Computer Science*, 2(420).

Solawetz, J.; and Francesco. 2024. YOLOv8. Roboflow Blog.

Tan, M.; and Le, Q. V. ???? EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. In Chaudhuri, K.; and Salakhutdinov, R., eds., *Proceedings of the 36th International Conference on Machine Learning, ICML 2019, 9-15 June 2019, Long Beach, California, USA*, volume 97 of *Proceedings of Machine Learning Research,*, 6105–6114. PMLR.

United Nations Environment Programme. 2024.

Wang, A.; Peng, T.; Cao, H.; Xu, Y.; Wei, X.; and Cui, B. 2022. TIA-YOLOv5: An improved YOLOv5 network for real-time detection of crop and weed in the field. *Front. Plant Sci.*, 13:1091655.

Yuan, L.; Chen, D.; Chen, Y.; Codella, N.; Dai, X.; Gao, J.; Hu, H.; Huang, X.; Li, B.; Li, C.; Liu, C.; Liu, M.; Liu, Z.; Lu, Y.; Shi, Y.; Wang, L.; Wang, J.; Xiao, B.; Xiao, Z.; Yang, J.; Zeng, M.; Zhou, L.; and Zhang, P. 2021. Florence: A New Foundation Model for Computer Vision. *CoRR*, abs/2111.11432.