A Deep Learning-based Approach for Vision-based Weeds Detection

Yan Wang*
Department of Architecture Engineering, Shijiazhuang College of Applied Technology, Shijiazhuang 050081, China

***Abstract*—Weed detection is an essential component of smart agriculture, and the use of remote sensing technologies has the potential to significantly improve weed management practices, reduce herbicide usage, and increase crop yields. This study proposed an approach to weed detection using computer vision and deep learning technologies. By utilizing remote sensing methods based on DL, this approach has the potential to optimize weed management strategies, minimize herbicide use, and enhance crop productivity. The weed detection algorithm is based on the Yolov8 framework, and a custom model is trained using images from popular datasets as well as the internet. To evaluate the model's effectiveness, it is tested on both validation and testing sets. Furthermore, the model's performance is assessed using images that are not included in the original dataset. As experimental results shown, the deep learning-based approach is a promising solution for weed detection in agriculture.**

Keywords*—*Smart agriculture; weed detection; remote sensing; deep learning; computer vision

I. Introduction

Smart agriculture, also known as precision agriculture [1], involves the use of technology to enhance and optimize crop production while minimizing environmental impact. This includes using sensors and monitoring systems to collect data on soil conditions, weather patterns, and plant growth, as well as utilizing machine learning algorithms and automation to improve efficiency and reduce waste. Smart agriculture also encompasses the use of drones, robotics, and other advanced technologies to perform tasks such as planting, watering, and harvesting. Overall, smart agriculture aims to increase yields, reduce resource usage, and promote sustainable farming practices [2].

Weeds are unwanted plants that compete with crops for resources such as nutrients, water, and sunlight, which can reduce the yield and quality of agricultural products [3]. Weeds can also serve as hosts for pests and diseases, which can further impact crop health and productivity. To manage weeds, farmers may use various methods such as mechanical cultivation, hand weeding, or chemical herbicides. However, the use of herbicides can have negative effects on the environment and human health, so there is growing interest in alternative weed management strategies such as integrated weed management, cover cropping, and crop rotation. Effective weed management is essential for maintaining healthy and productive agricultural systems while minimizing negative impacts on the environment and human health [4, 5].

Weed detection is a critical aspect of precision agriculture, as it allows farmers to identify and manage weeds more effectively [6]. There are several methods for detecting weeds in agricultural fields, including visual inspections, manual sampling, and remote sensing technologies. Visual inspections involve physically observing the crop and looking for signs of weed growth. This method can be time-consuming and labor-intensive, but it can be useful for identifying small infestations or for crops with lower weed densities. Manual sampling involves collecting samples of soil or plant material from different locations in the field and analyzing them for the presence of weeds [7]. This method is more accurate than visual inspections but can still be time-consuming and requires trained personnel. Remote sensing technologies, such as satellites, drones, or ground-based sensors, can provide rapid and accurate weed detection across large areas [8, 9]. These technologies use various sensors such as multispectral or hyperspectral cameras, which can detect differences in plant color, reflectance, or texture, to identify and map weeds [10]. Machine learning algorithms can then be used to analyze the data and develop weed management strategies [11].

Deep learning is a subset of machine learning that uses artificial neural networks to automatically learn and identify patterns in data [11, 12]. In the context of weed detection, deep learning algorithms can be trained on large datasets of images to recognize and classify different weed species. This involves using convolutional neural networks (CNNs) to extract features from the images and then using a combination of fully connected layers and softmax classifiers to classify the images [13]. Deep learning-based weed detection systems have shown high accuracy rates in detecting and classifying weeds, even in complex agricultural environments [14, 15]. These systems have the potential to revolutionize weed management practices by providing farmers with rapid and accurate information about weed infestations, allowing for more targeted and effective weed control strategies.

In this study, a deep learning-based method is proposed for weed detection. In DL based which is in remote sensing technologies it has the potential to significantly improve weed management practices, reduce herbicide usage, and increase crop yields. In order to detect weeds, a Yolo based algorithm is developed to detect the weeds. For this detection, a model is trained using collected images from internet and other popular dataset. The generated model is evaluated and test using associated validation and testing sets. Finally, the model is tested with images outside of our dataset to make sure the performance of the method is effective.

The main research contributions of this study are as follows:

1) The study introduces a novel deep learning-based method for weed detection in remote sensing technologies, offering the potential to enhance weed management practices and contribute to reductions in herbicide usage, ultimately leading to increased crop yields.

2) A YOLO-based algorithm is developed as part of the research, providing an effective and efficient means of detecting weeds, thereby contributing to the advancement of automated weed detection systems.

3) The research contributes by presenting a meticulously trained model, utilizing a diverse set of images collected from the internet and other popular datasets, and evaluates its performance not only on associated validation and testing sets but also on external images, ensuring the method's effectiveness beyond the initial dataset.

II. Related Works

In research [12], custom lightweight deep learning models are suggested for detecting weeds in soybean crops. The models were trained using a dataset of images depicting soybean crops with varying weed types. The findings demonstrate that the proposed models outperform conventional machine learning algorithms in terms of speed, memory usage, and accuracy. The authors suggest that the custom lightweight deep learning models can efficiently detect weeds in soybean crops, which can result in improved crop management and reduced usage of herbicides.

Peng et al, [16] presented an enhanced RetinaNet network for detecting weeds in paddy fields. The proposed network employs residual connections and feature pyramid network (FPN) to improve the accuracy of weed detection. The dataset used in this study consists of images of paddy fields containing different types of weeds, which were used for both training and testing the network. The findings of the study indicate that the proposed network outperforms existing methods and is effective in detecting weeds in paddy fields. The authors suggest that the improved RetinaNet network has the potential to aid in weed management and reduce the need for herbicides in agriculture.

Haq et al, [17] developed an automated weed detection system that relies on CNNs and UAV imagery. The proposed system captures aerial images of crop fields using UAVs and employs CNNs to differentiate between crops and weeds. The CNNs are trained using a dataset of images containing both crops and weeds. The study reveals that the system can accurately detect weeds in a timely manner. Their experimental results show using UAVs and CNNs for weed detection can lead to better weed management in agriculture, enhancing efficiency and accuracy.

The authors in study [18] presented an enhanced version of the YOLO v4 algorithm for detecting weeds in images of carrot fields. The proposed algorithm leverages data augmentation and transfer learning techniques to improve the performance of the YOLO v4 model. The authors collected a dataset of images of carrot fields with and without weeds, and the algorithm was trained and evaluated on this dataset. The findings indicate that the improved YOLO v4 algorithm surpasses the traditional YOLO v4 and other advanced algorithms. The authors suggest that the algorithm can effectively identify weeds in carrot fields, contributing to weed management and minimizing the use of herbicides. The study highlights the potential of deep learning methods in agriculture for weed detection.

Alam [5] proposed a machine-learning based system for real-time crop/weed detection and classification, facilitating variable-rate spraying in precision agriculture. The system employs a CNN to detect and differentiate crops and weeds based on their visual features. The study reveals that the proposed system accurately identifies crops and weeds, which can improve targeted spraying and minimize herbicide usage. This study showed that the machine-learning based approach can enhance precision agriculture and promote sustainable crop management practices.

III. Methodology

A. Yolov8 Algorithm

The YOLOv8 is an efficient object detection model that was introduced in early of 2023 [19]. It is an improvement over the previous versions of YOLO, which are known for their speed and accuracy in object detection. YOLOv8 is designed to be more accurate than its predecessors while still maintaining real-time performance. The YOLOv8 architecture is composed of several components, including a backbone network, a neck network, and a head network. Fig. 1 shows the architecture of YOLOv8 network.

The backbone network is responsible for extracting features from the input image, while the neck network and the head network are responsible for detecting objects and generating bounding boxes. It uses several equations and formulas in its implementation, including:

1) Sigmoid function: YOLOv8 uses a sigmoid function to transform the predicted outputs into probabilities. The sigmoid function is defined as follows:

$$\text{sigmoid}(x) = 1 / (1 + e^{-x}) \tag{1}$$

where, x is the input to the function.

2) Intersection over Union (IoU): IoU is used to measure the overlap between two bounding boxes. It is defined as the ratio of the area of the intersection of the two bounding boxes to the area of their union.

$$\text{IoU}(A, B) = (\text{area of intersection between A and B}) / (\text{area of union between A and B}) \tag{2}$$

3) Anchor boxes: YOLOv8 uses anchor boxes to predict the location and size of objects in the image. Anchor boxes are fixed bounding boxes of different sizes and aspect ratios that are placed at various locations in the image.

4) Loss function: The loss function used in YOLOv8 is a combination of three different losses: the localization loss, the confidence loss, and the class loss. The localization loss measures the difference between the predicted and ground-truth bounding box coordinates. The confidence loss measures the difference between the predicted and ground-truth objectness scores. The class loss measures the difference between the predicted and ground-truth class probabilities.

5) YOLOv8 output tensor: The output of YOLOv8 is a tensor that contains predictions for each anchor box. Each anchor box has a corresponding set of predicted values, which include the class probabilities, objectness score, and bounding box coordinates. The tensor is typically represented as follows:

[batch_size, grid_size, grid_size, num_anchors, num_classes + 5] (3)

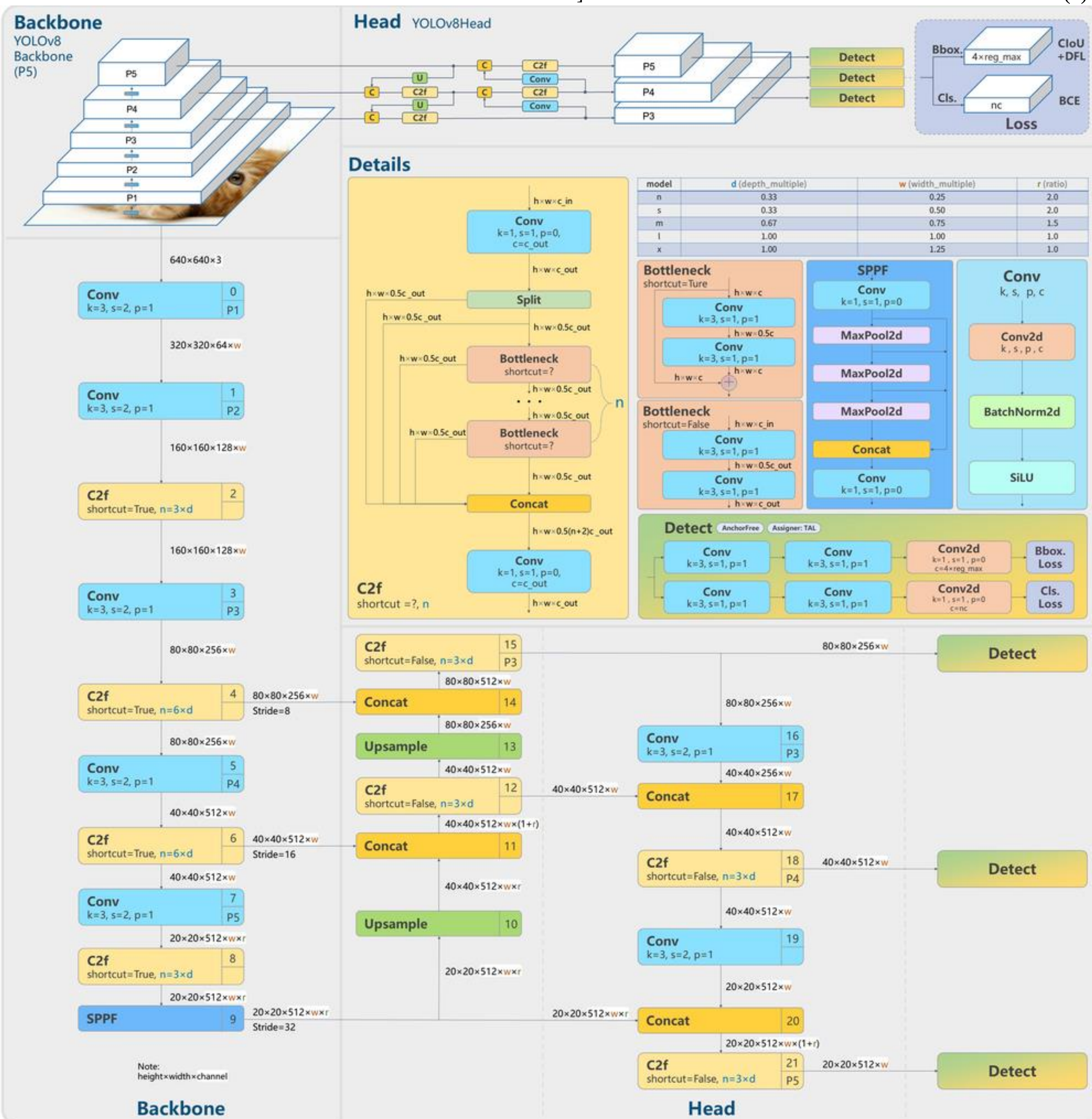

Fig. 1. Architecture of YOLOv8 network [20].

B. Dataset

In this study, a dataset is used from internet resource. The dataset includes images taken from a public dataset in Roboflow. We have totally 4239 images in the dataset. Among these images, augmentation process is performed for extending the dataset. The structure for training task from this dataset, 87% or 3700 images for training set, 9% or 359 images for validation set, and 4% or 180 images for testing set are organized. Some images from the dataset are shown in Fig. 2.

C. Google Colab

To conduct our experiments, we utilized Google's Colab research platform, which offers access to high-performance GPUs at no cost. We conducted all of our training and testing on a 12GB NVIDIA Tesla T4 GPU, which is described in more detail in Fig. 3. Our models were trained with a maximum of 2500 iterations, a batch size of four images, and an image size of 640.

D. Comparison of Yolo Models

This section presents a companion of different models of Yolo networks, the purpose of this comparison is to justify why Yolov8 is selected in this study. Based on published performance analysis of different Yolo based models [20], this investigation is conducted. For this investigation, we can analyze the model's graph where the X-axis represents the mean average precision (mAP) percentage, and the Y-axis represents the number of parameters in each YOLO-based model. The graph displays curves corresponding to different YOLO model versions: YOLOv8, YOLOv7, YOLOv6, and YOLOv5. Fig. 3 shows this graph.

Fig. 2. Sample images of the dataset [20].

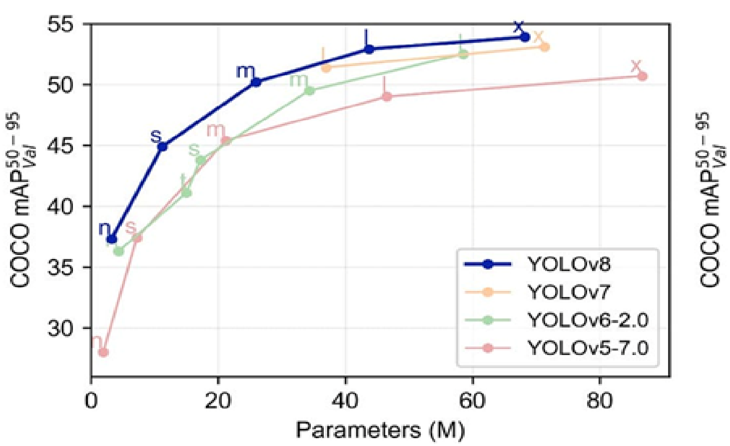

Fig. 3. Comparison of Yolo models in terms of mAP [20].

The graph depicts a comparison between the mAP percentages and the number of parameters for each YOLO-based model. The curves demonstrate the trade-off between the mAP performance and the complexity of the models, represented by the number of parameters. YOLOv8, having the best mAP performance among the models, exhibits a curve that consistently outperforms the other models in terms of mAP percentage. This indicates that YOLOv8 achieves higher accuracy without excessively increasing model complexity, making it a more efficient and scalable choice. Therefore, based on the assumptions provided, the graph illustrates that YOLOv8 surpasses the other models in terms of mAP performance while maintaining a reasonable number of parameters. This makes YOLOv8 the preferred option among the YOLO-based models considered in the graph, as it offers superior accuracy without excessive model complexity.

Moreover, we can analyze the graph where the X-axis represents the average precision (AP) percentage, and the Y-axis represents the performance of PyTorch FP16 running on the RTX 3080 platform. The performance measurements are conducted on the COCO dataset. The graph includes curves corresponding to different YOLO-based model versions: YOLOv8, YOLOv8-seg, YOLOv7, YOLOv6, YOLOv6, and YOLOv5.

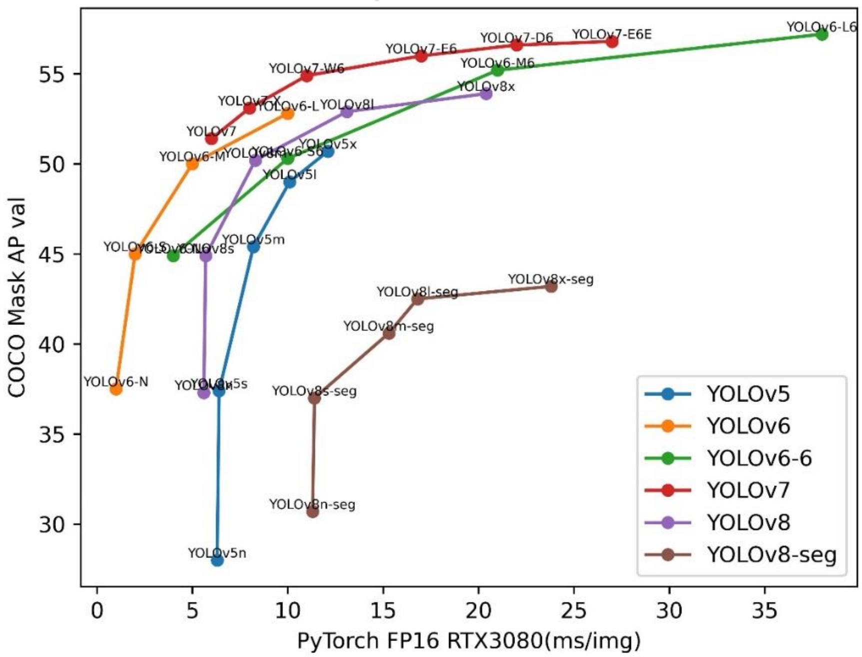

Fig. 4. Comparison of performance for Yolo models in terms of AP [20].

As shown in Fig. 4, the graph presents a comparison between the AP percentages and the performance of PyTorch FP16 running on the RTX 3080 platform for each YOLO-based model. The curves showcase the relationship between AP performance and the computational efficiency of the models. YOLOv8-seg, being the model with the best AP performance according to the assumption, exhibits a curve that consistently outperforms the other models in terms of AP percentage. This indicates that YOLOv8-seg achieves higher accuracy in object detection on the COCO dataset compared to the other versions.

The YOLOv8-seg's superiority is justified not only in terms of AP but also in relation to the computational efficiency represented by the PyTorch FP16 performance on the RTX 3080 platform. Despite its superior AP performance, YOLOv8-seg manages to maintain efficient performance on the given hardware platform, suggesting that it strikes a balance between accuracy and computational efficiency. Therefore, based on the assumptions provided, the graph illustrates that YOLOv8-seg outperforms the other models in terms of AP performance on the COCO dataset while maintaining efficient performance on the specified hardware platform. This makes YOLOv8-seg the preferred choice among the YOLO-based models considered in the graph, as it offers higher accuracy without compromising computational efficiency.

IV. Experimental Results

This section presents experimental results and discuss about performance evaluation in details. Firstly, experimental result is presented from the trained model using above details, and then performance evaluation is discussed. Fig. 5 shows some experimental results.

Performance evaluation is an essential step in the development of object detection models, including the YOLO object detection algorithm. Model evaluation helps assess the performance of a model and determine whether it is meeting the desired accuracy criteria. Popular performance metrics used for model evaluation in object detection tasks are precision, recall, and mAP (mean Average Precision). Fig. 6 shows the performance results.

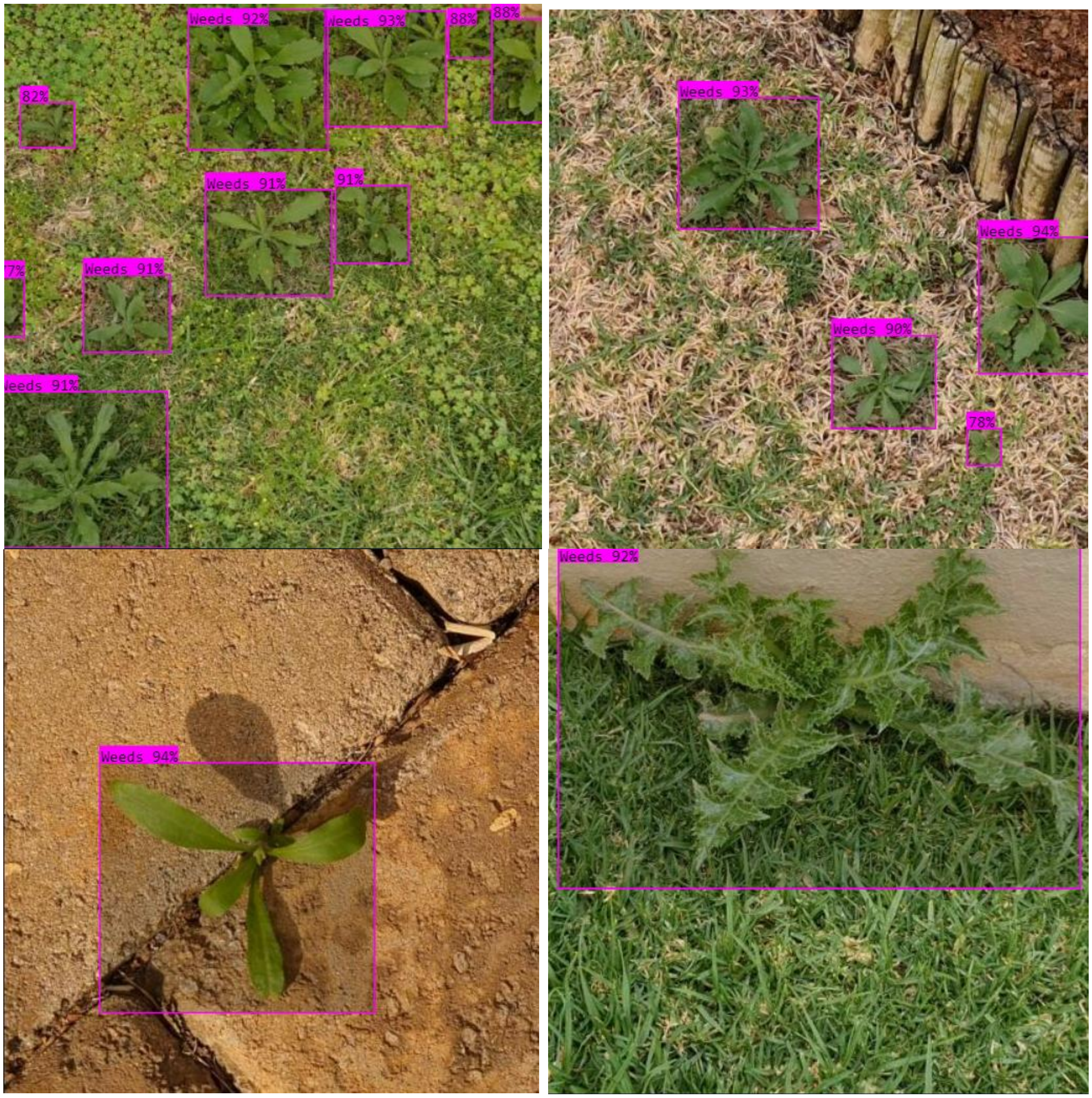

Fig. 5. The result of our experiments.

To present the statistical reporting and data presentation in evaluating the YOLOv8-based model for weed detection, this study provides a detailed breakdown of precision, recall, and mAP values. Instead of a single mAP value, presenting precision-recall curves at different confidence thresholds can offer a nuanced understanding of the model's performance across a range of decision-making points. This not only adds depth to the analysis but also provides insights into the trade-off between precision and recall, aiding in decision-making for real-world applications. Additionally, including a confusion matrix or a similar visual representation would offer a more granular view of the model's strengths and weaknesses, particularly in terms of false positives and false negatives.

As shown in Fig. 6, precision, recall, and mAP are essential metrics used to assess the performance of a model, including its effectiveness in weed detection:

Precision: Precision measures how well the model predicts true positive instances while minimizing false positives. High precision indicates that the model is accurate in identifying weeds and has fewer false alarms.

Recall: Recall measures the model's ability to correctly identify all positive instances, or in this case, accurately detecting weeds. High recall suggests that the model is effective at capturing most of the weeds present in the dataset.

mAP: mAP provides a comprehensive evaluation of the model's performance by considering both precision and recall at various thresholds. A higher mAP indicates a more accurate and effective model for weed detection.

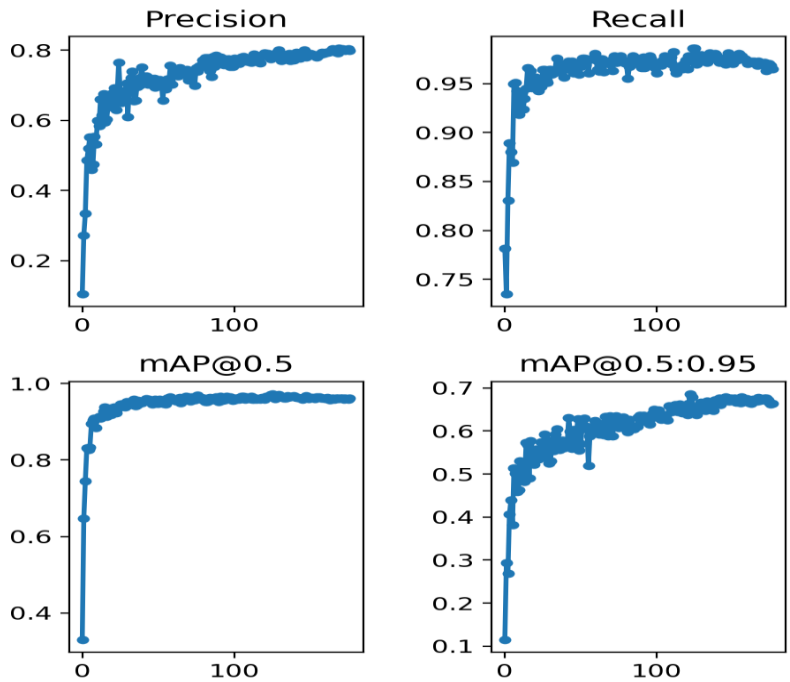

Fig. 6. The performance results.

By considering the precision, recall, and mAP metrics, we can assess the effectiveness of the generated YOLOv8 model for weed detection. If the precision curve demonstrates high precision values across different thresholds, it indicates that the model reliably detects weeds while minimizing false positives. A steep rise in precision suggests the model is precise at differentiating between weeds and non-weed instances. Similarly, if the recall curve shows high recall values, it implies that the model successfully captures a significant number of weed instances, reducing the chances of missing any weeds. Lastly, a high mAP indicates a balance between precision and recall, indicating that the model achieves both accurate weed detection and minimizes false positives.

Therefore, by evaluating the precision, recall, and mAP metrics and ensuring high values for all these measures, we found that the generated YOLOv8 model is effective for accurate weed detection.

V. Conclusion

In this research, a deep learning approach for vision-based weeds detection in agriculture is proposed. The algorithm used for weed detection is built on the Yolov8 framework, and a customized model is created by training it on images from popular datasets as well as the internet. To assess the effectiveness of the model, it is tested on both validation and testing datasets, and its performance is evaluated using images that are not part of the original dataset. The experimental findings demonstrate that the deep learning-based approach is a promising solution for detecting weeds in agriculture. However, in this research for limitation addressing purpose, the diversity of the training data sources used for creating the YOLOv8-based weed detection model. While the model is trained on images from popular datasets and the internet, the potential presence of biases in these sources may affect the model's generalizability to a broader range of real-world agricultural scenarios. The use of internet-collected images might introduce variations in terms of lighting conditions, field types, and weed species that are not fully represented in the training dataset. Consequently, the model's performance might be over-optimized for the specific characteristics of the training data, limiting its effectiveness in more diverse and unpredictable agricultural environments. To address this limitation, future studies should focus on systematically expanding the diversity of the training dataset to ensure the model's robustness across a broader spectrum of agricultural conditions. This could involve incorporating images from geographically diverse locations, different seasons, and various agricultural practices. Additionally, efforts should be made to include images that capture the inherent variability in weed species and growth stages, ensuring that the model can accurately detect weeds under a wide range of circumstances. By addressing this limitation, researchers can enhance the model's applicability and reliability in real-world agricultural settings, ultimately contributing to the successful deployment of the proposed deep learning approach on a larger scale.

References

[1] I. Cisternas, I. Velásquez, A. Caro, and A. Rodríguez, "Systematic literature review of implementations of precision agriculture," Computers and Electronics in Agriculture, vol. 176, p. 105626, 2020.

[2] D. C. Tsouros, S. Bibi, and P. G. Sarigiannidis, "A review on UAV-based applications for precision agriculture," Information, vol. 10, no. 11, p. 349, 2019.

[3] S. Shanmugam, E. Assunção, R. Mesquita, A. Veiros, and P. D. Gaspar, "Automated weed detection systems: A review," KnE Engineering, pp. 271–284-271–284, 2020.

[4] S. Kulkarni, S. Angadi, and V. Belagavi, "IoT based weed detection using image processing and CNN," International Journal of Engineering Applied Sciences and Technology, vol. 4, no. 3, pp. 606-609, 2019.

[5] M. Alam, M. S. Alam, M. Roman, M. Tufail, M. U. Khan, and M. T. Khan, "Real-time machine-learning based crop/weed detection and classification for variable-rate spraying in precision agriculture," in 2020 7th International Conference on Electrical and Electronics Engineering (ICEEE), 2020: IEEE, pp. 273-280.

[6] B. Liu and R. Bruch, "Weed detection for selective spraying: a review," Current Robotics Reports, vol. 1, pp. 19-26, 2020.

[7] N. Islam et al., "Early weed detection using image processing and machine learning techniques in an Australian chilli farm," Agriculture, vol. 11, no. 5, p. 387, 2021.

[8] A. Wang, W. Zhang, and X. Wei, "A review on weed detection using ground-based machine vision and image processing techniques," Computers and electronics in agriculture, vol. 158, pp. 226-240, 2019.

[9] M. D. Bah, E. Dericquebourg, A. Hafiane, and R. Canals, "Deep learning based classification system for identifying weeds using high-resolution UAV imagery," in Intelligent Computing: Proceedings of the 2018 Computing Conference, Volume 2, 2019: Springer, pp. 176-187.

[10] Z. Wu, Y. Chen, B. Zhao, X. Kang, and Y. Ding, "Review of weed detection methods based on computer vision," Sensors, vol. 21, no. 11, p. 3647, 2021.

[11] A. M. Hasan, F. Sohel, D. Diepeveen, H. Laga, and M. G. Jones, "A survey of deep learning techniques for weed detection from images," Computers and Electronics in Agriculture, vol. 184, p. 106067, 2021.

[12] N. Razfar, J. True, R. Bassiouny, V. Venkatesh, and R. Kashef, "Weed detection in soybean crops using custom lightweight deep learning models," Journal of Agriculture and Food Research, vol. 8, p. 100308, 2022.

[13] D. Patel, M. Gandhi, H. Shankaranarayanan, and A. D. Darji, "Design of an Autonomous Agriculture Robot for Real-Time Weed Detection Using CNN," in Advances in VLSI and Embedded Systems: Select Proceedings of AVES 2021: Springer, 2022, pp. 141-161.

[14] C. T. Selvi, R. S. Subramanian, and R. Ramachandran, "Weed Detection in Agricultural fields using Deep Learning Process," in 2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS), 2021, vol. 1: IEEE, pp. 1470-1473.

[15] A. Nasiri, M. Omid, A. Taheri-Garavand, and A. Jafari, "Deep learning-based precision agriculture through weed recognition in sugar beet fields," Sustainable Computing: Informatics and Systems, vol. 35, p. 100759, 2022.

[16] H. Peng, Z. Li, Z. Zhou, and Y. Shao, "Weed detection in paddy field using an improved RetinaNet network," Computers and Electronics in Agriculture, vol. 199, p. 107179, 2022.

[17] M. A. Haq, "CNN Based Automated Weed Detection System Using UAV Imagery," Comput. Syst. Sci. Eng., vol. 42, no. 2, pp. 837-849, 2022.

[18] B. Ying, Y. Xu, S. Zhang, Y. Shi, and L. Liu, "Weed Detection in Images of Carrot Fields Based on Improved YOLO v4," Traitement du Signal, vol. 38, no. 2, 2021.

[19] C. H. Kang and S. Y. Kim, "Real-time object detection and segmentation technology: an analysis of the YOLO algorithm," JMST Advances, pp. 1-8, 2023.

[20] K. Chen et al., "MMDetection: Open mmlab detection toolbox and benchmark," arXiv preprint arXiv:1906.07155, 2019. [Online]. Available: https://github.com/open-mmlab/mmyolo/tree/dev/configs/yolov8.