Indian Journal of Weed Science (2024) **56**(2): 194–199
http://dx.doi.org/10.5958/0974-8164.2024.00031.X

RESEARCH ARTICLE

Deep learning-based weed detection in sesame crops using modified YOLOv5 model

Sandip Sonawane* and Nitin N. Patil

Received: 19 October 2023 | Revised: 22 June 2024 | Accepted: 26 June 2024

ABSTRACT

In agriculture, the weed plant identification is a challenging task as it allows farmers to accurately recognize and remove the same plants from their field. In India, conventional methods for detecting and removing weeds require considerable manual labor and skill, resulting in a time-consuming and costly process. With recent advancements in machine learning and computer vision, automated weed detection systems have become more prevalent. We worked an innovative method for crop-weed classification and weed detection that utilizes a Convolutional Neural Network (CNN) to differentiate images of plant into either weed or non-weed categories. The techniques we introduced were developed using an extensive dataset containing 1300 images of sesame (*Sesamum indicum L.*) crops cultivated in farmlands of China. The proposed approach evaluated on a dataset available on the Roboflow platform. We used ResNet50 architecture for image classification and Faster-RCNN and YOLO (You Only Look Once) for object detection. The YOLOv5 model's performance was measured by utilizing Precision (P), Recall (R), and the mean Average Precision (mAP) as performance evaluation metrics. The proposed modified YOLOv5 model achieved the best overall performance within the 'Weeds' validation subset resulting in a P (80.7), R (81.1), and mAP (86.4). This approach is suitable for bermudagrass, crabgrass and pigweed species of weeds in sesame field. The proposed approach has several practical applications in agriculture, including weed management, crop yield optimization, and environmental sustainability. Furthermore, it has potential use when integrated with other precision farming equipment, making it a cost-effective solution for farmers. We concluded the efficacy of employing deep learning methods for the detection of weed plants and suggest that it has the potential to revolutionize modern agriculture.

Keywords: Convolutional Neural Network, Deep learning, Sesame, Weed detection, YOLO model

INTRODUCTION

One of the most important factors of overall crop productivity is weed. Weed competes with crop for assets like water, space, nutrients and light. This competition of weeds with crop plants decreases yields and degrades production quality. Weeds can also host diseases and pests, which can further harm the crops (Singh and Gupta 2022). Furthermore, weeds lower the quality of crops by contaminating them with weed seeds or by making them tough to harvest. The cost of controlling weeds can also be important, as it often needs the herbicides usage or manual work to eliminate them. The present research indicates that the presence of weeds can lead to a reduction in overall productivity ranging from 10% to 90%, depending on the extent of the infestation and the crop variety (Nurudeen *et al.* 2024). To mitigate their influence on crop yields and maintain the sustainability of agricultural systems, it is crucial to effectively manage weeds.

Department of Computer Engineering, R.C. Patel Institute of Technology, Shirpur, Maharashtra 425405, India

* Corresponding author email: sandip.sonawane@rcpit.ac.in

Weed management plays a crucial role in all crops. Managing weeds is an important task of forestry practices and agriculture (Elhoseny *et al.* 2023), as an uncontrolled growth of weeds can considerably reduce crop production and overall quality (Alotaiby *et al.* 2022). Traditionally weeds are being managed by several methods such as hand weeding and overall chemical spraying but these methods have proven as time-consuming, costly, labor-intensive, and environmentally dangerous (Shanmugam *et al.* 2021). Hence, it is necessary to develop a precision weed management approach that accurately detects weeds and controls them. The overall objective of the weed plant detection method is to identify and localize the presence of weeds accurately in an image or video stream. Because of the high variation in color, texture, appearance and shape of weed species, along with their complexity, this is a particularly difficult task.

With the growth of machine learning algorithms and computer vision over the recent years, it is now possible to develop automated systems for detecting and identifying weeds in agricultural fields. By providing timely and precise information about the location and distribution of weeds, these systems can assist farmers in accurately detecting the weed species and subsequently controlling them. This approach leads to minimizing the use of herbicides, and lowering production costs, increasing crop yields and ensuring the sustainability of agricultural systems. Numerous traditional machine-learning approaches, relying on image processing methods using classifiers such as K-Nearest Neighbor (KNN), Random Forest (RF), Support Vector Machine (SVM) and Decision Tree have been applied. These approaches use various feature extraction techniques, including shape, texture and color analysis, for the detection and categorization of weeds. Furthermore, with the help of deep learning algorithms, these systems can learn from large datasets and gradually improve their performance (Pallottino *et al.* 2022).

In this work, we explored deep-learning techniques for detection and differentiation classification of crop-weeds species. Specifically, we use a CNN to discover weeds in high-resolution imageries.

MATERIALS AND METHODS

The suggested approach involves several steps represented in pictorial form as shown in **Figure 1**. We conducted our experiments on a publicly accessible dataset that is available on the Roboflow platform. The study utilized a dataset of weed and sesame crop images cultivated during the spring season from row-based farmlands located in Nanning, Guangxi, China. The images were acquired during kharif season and under cloudy weather conditions. The variety of sesame used in the study was '*Yuzhi 11*', which has a growth duration of approximately 90-100 days (Jiqing *et al.* 2022). The dataset comprises 1300 images, each having a resolution of 4000 by 3000 pixels. Images were captured at various growth stages of the sesame crop, including germination, vegetative, flowering, and pod-filling stages. The crop was sown using the line sowing method, with a crop geometry of 30 cm row-to-row and 10 cm plant-to-plant distance. The average plant population observed in the field was approximately 330,000 plants per hectare. The weeds observed in sesame crop includes Bermudagrass (*Cynodon dactylon*), Crabgrass (*Digitaria* spp.), Pigweeds (*Amaranthus* spp.) *etc*. We converted all

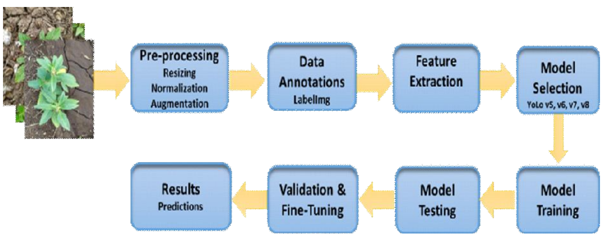

Figure 1. Process flow of weed detection and classification system

images to 640 x 640 x 3 pixel size. To reduce noise and enhance the image contrastness, we pre-processed the input image.

The dataset images were manually annotated to discriminate between sesame crops and weeds using the LabelImg tool. All image were meticulously reviewed, and regions containing sesame crops and weeds were labeled to create ground truth data with precise locations and boundaries. Based on the shape, color and texture features, the ground truth data distinguishes between crop and weed. This approach aligns with supervised learning, where the model is trained on input-output pairs to learn the mapping from inputs to the desired outputs. Deep learning benefitted from automatic feature extraction via layers of Convolutional Neural Networks (CNNs).

The YOLOv5 object detection model was selected and trained on an extensive dataset to enhance the model's accuracy. The proposed work evaluates performance in terms of accuracy, precision, and recall, demonstrating their effectiveness in detecting weeds and differentiating them from sesame crop. To further improve the performance of the proposed models, hyperparameter fine-tuning was conducted. The dataset was divided into testing (10%), validation (20%), and training (70%) subsets. We used 900 images for training, 137 for testing and 263 images for validation purposes.

Image classification and object detection models

The proposed work utilized ResNet50 model for image classification and Faster-RCNN, YOLOv5 model for object detection. Proposed model utilized Transfer Learning (TL) methods on pre-trained models ImageNet, VGG16 and ResNet50 leveraging both the Keras and PyTorch frameworks.

A notable example of a two-stage object detection framework is Faster R-CNN. Initially, it generates region proposals through a Region Proposal Network (RPN), followed by refining these proposals for the ultimate purpose of object detection and classification. YOLO stands as a renowned single-stage object detection algorithm that processes the entire image in one forward pass through a neural network. Both Faster-RCNN and YOLOv5 primarily uses the PyTorch framework. Faster R-CNN has been implemented in other deep learning frameworks as well, such as TensorFlow.

The YOLO architecture has consistently been a widely accepted model for object recognition among deep learning professionals. In June 2020, Ultralytics introduced the state-of-the-art object detection model YOLOv5. It represents an improvement over the YOLOv4 framework, which is renowned for its outstanding accuracy and ability to operate in real-time.

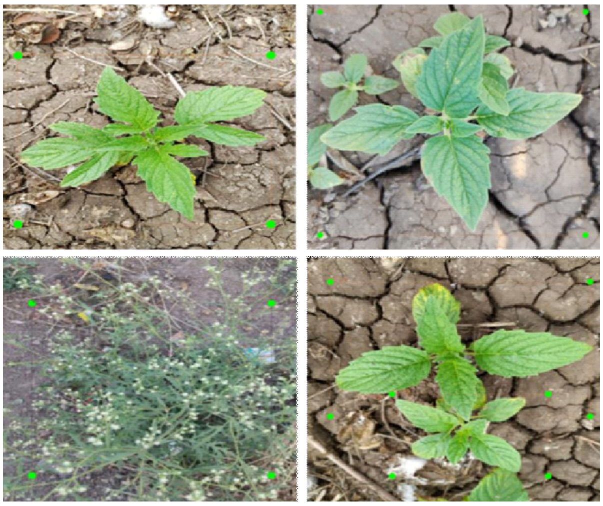

Figure 2. Sample images of sesame crop and weed in the dataset

YOLOv5 features a single-layer object detection network with a CSPDarknet53 feature extractor as the backbone. The model structure incorporates several innovative elements like Spatial Pyramid Pooling (SPP), PAN, and BiFPN, all contributing to enhancing the effectiveness and accurateness of the trained model. The YOLOv5 model has attained leading-edge performance on various benchmark object detection datasets, including Pascal Visual Object Classes (VOC) and Microsoft Common Objects in Context (COCO) (Sportelli *et al.* 2023).

As shown in **Figure 3,** the YOLOv5 architecture relies on a fusion of a cross-stage partial network (CSPNet) and the Darknet, which serves as its foundational framework (C. Y. Wang *et al.* 2016). To fulfill the requirements of the YOLO algorithm, images were annotated using LabelImg tool. LabelImg stores annotations in a variety of formats such as XML, JSON, CSV, or text format (López-Correa *et al.* 2022). In YOLO, bounding box information was stored in text file format following a particular syntax. The rectangular (bounding) box were recorded on a separate line and included five numerical values. The initial number indicated the label of class, while the second and third numbers denoted the x and y coordinates of the top-left corner of the bounding box. The 4th and 5th numbers denoted the bounding box's width and height (Aanis Ahmad *et al.* 2021).

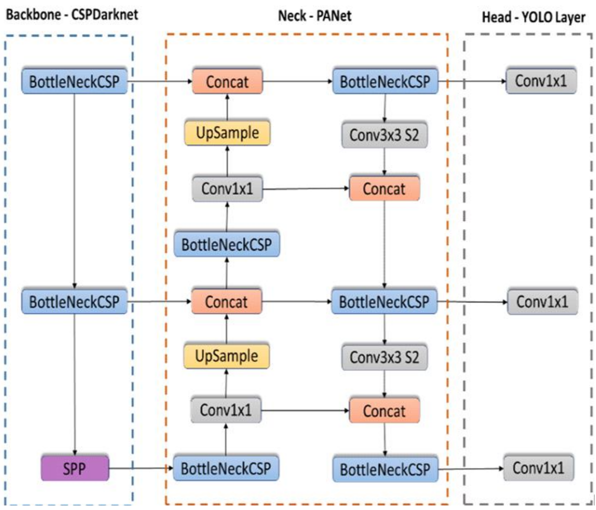

Figure 3. Architecture of YOLOv5

The annotation process were applied to all plants within the images, and all the relevant information of annotations of bounding box were stored on a single line of text.

We implemented our proposed methods in Python using TensorFlow and OpenCV libraries. We used a workstation with an Intel Core i3 CPU, 8GB of RAM, and a Google Colab to train and evaluate our models.

We have standardized the following hyper-parameters across all experimental configurations for YOLOv5.

- Training epochs: 100 and 150
- Solver type/Optimizer: SGD (Stochastic Gradient Descent)
- Input image size: 640
- Momentum: 0.937
- Batch size: 4
- Learning rate (LR) policy: Exponential decal
- Weight decay: 0.0004
- Base learning rate: 0.0001
- AutoAnchors: 3.51 anchors/target

To assess the efficiency of the suggested approach, we employed the evaluation metrics: accuracy in the form of mAP (mean Average Precision), recall and precision.

Evaluation metrics

In this proposed work, we used precision, recall and mAP as the evaluation criteria for training the detection model (Jialin *et al.* 2019).

Precision is defined as the ratio of correctly identified weeds of a specific species among the expected weeds. Recall represented the percentage of correctly predicted targets within a weed class within the sample. The following is the formula (Tushar *et al.* 2023):

$$P = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (1)$$

$$R = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (2)$$

In this context, True Positive (TP) denotes the samples count correctly categorized as positive samples, False Positive (FP) signifies the count of erroneously classified positive samples, and FN represented for incorrectly classified negative samples count.

The mAP represents the average of the individual average precisions calculated for all categories within the dataset. This is computed by dividing the summation of average precisions for all categories by the total number of categories:

$$mAP = \frac{\sum AP}{n} \tag{3}$$

RESULTS AND DISCUSSION

The training performance of ResNet50 and Faster-RCNN is given below in **Table 1**. The proposed work achieved accuracy of 81.6% with Faster-RCNN and 79.4% with ReesNet50. The training results of the YOLOv5s model is given below. The **Table 2** provides a comparison between two configurations of the YOLOv5s model in terms of training parameters and performance metrics. The second configuration, which were trained for more epochs, generally outperforms the first configuration in terms all evaluation metrics. From the results, it seems that YOLOv5 outperformed in terms of accuracy.

Table 1. Training results of ResNet-50 and Faster-RCNN

Parameter	ResNet50	Faster-RCNN
No. of training Steps	25,000	25,000
Training time	7 Hrs. (1 sec/step)	7 Hrs. (1 sec/step)
Loss	0.0067	0.0843
Learning Rate	0.0165	0.0185
Accuracy	79.4%	81.6%

Table 2. Training results of YOLOv5s

Parameters	YOLOv5s	
Epochs	100	150
Training time	0.737 hours	1.036 Hours
Precision	76.70	80.70
Recall	78.00	81.11
Accuracy (mAP@0.5)	82.90	86.40
mAP@0.5-0.95	49.70	44.80

Figure 4 given below presented inferencing / detection results (predictions) of YOLOv5 with a confidence level (IoU) of 0.5 on validation dataset visualize the results.

These examination outcomes were displayed using an IoU threshold with 0.5. (In this illustration, the class label, which can be either “weed” or “crop,” is situated to the left corner of the bounding box, while the precision score for that class is positioned to the right.). The curves of YOLOv5 are shown below in figure 5. Here, the precision-recall curve of YOLOv5 revealed that the weed class achieved a slightly superior average precision (AP) score of

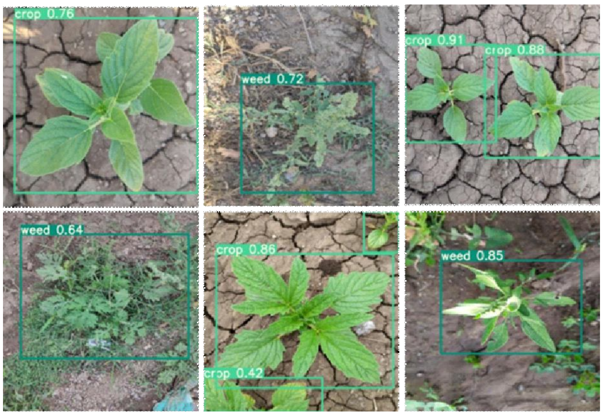

Figure 4. Detection results using YOLOv5 with a bounding box

0.877 compared to the crop class. The solid blue line fig mAP at an 0.5 IoU, calculated on test dataset. The individual average precision (AP) scores for each class and the overall algorithm’s mAP were reflective of the area beneath their respective curves on the graph.

In **Figure 5 (a)** A curve that represents the relationship between precision and confidence. From this curve it is cleared that, precision is maximum with a confidence level of 0.877, suggesting a significant proportion of true positive results across all classes. (b) A curve that illustrates the connection between recall and confidence. The recall-confidence curve analysis provides insights into prediction performance, exhibiting a progressive drop in recall values as confidence levels increase. (c) A curve that shows the interplay between precision and recall. Here class 1(weed) achieves a slightly superior average precision of 0.877 than class 0. (d) A curve displaying the connection between F1 score and confidence. The average F1 score reached to 0.81 with a confidence interval of 0.374.

During the fine-tuning of hyper parameters, the metrics and losses are still improving and this is depicted in **Figure 6**. It’s evident that the box loss, obj loss and cls loss parameters in both the datasets (training and validation) of the trained model consistently decreased. Simlutaneously, the AP with mAP@0.5 consistently improved. YOLOv5 achieved an mAP@0.5 score near to 0.9, signifying superior training outcomes when using the sesame dataset. The results were recorded in the results.csv file after each epoch and are subsequently visualized as results.png upon completing the process of training. Additionally, we can create plots manually using any results.csv file. After training our model, we achieved less loss value in both the validation and training as given in the following statistics in **Table 3**.

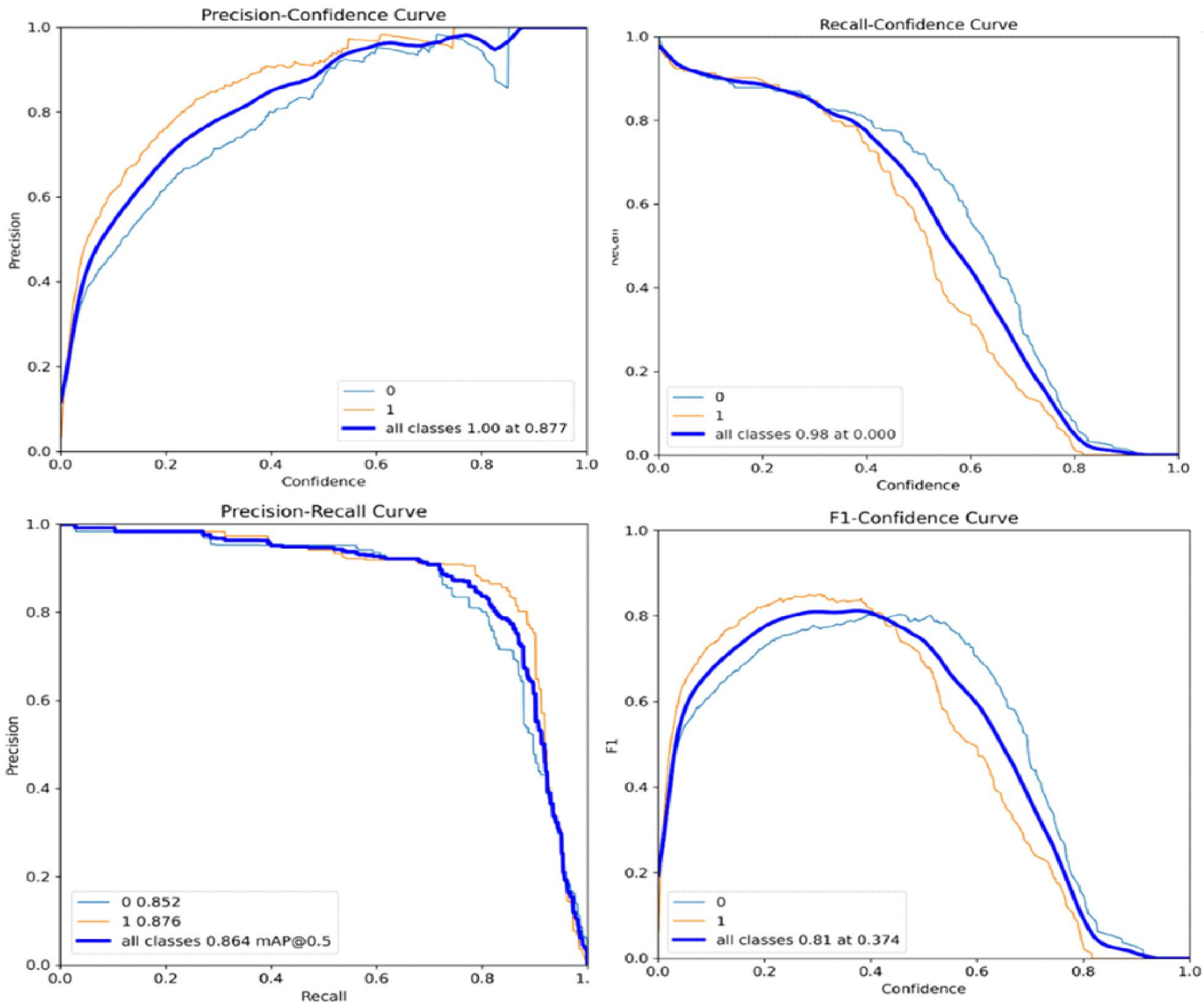

Figure 5. Graphical representation of performance parameters for training via YOLOv5

Table 3. Training and validation loss

Phase	box_loss	obj_loss	cls_loss
Training	0.0288	0.0218	0.0035
Validation	0.0215	0.0086	0.0043

Our suggested approach attained 86.40% accuracy, 80.70% precision, and 81.11% recall in detection of weed species in the dataset. YOLO method outperformed both other methods ResNet and Faster RCNN regarding evaluation metrics precision, recall, accuracy and inferencing speed, demonstrating the superiority of our weed detection approach.

Similar approach was implemented by (Chen *et al*. 2022) who reported that enhancing YOLOv4 model with an attention mechanism and adaptive spatial feature fusion achieved high performance in terms of various metrics.

Dhruw *et al* 2023 in their study proposed three popular object detection algorithms for detecting weeds in soybean plantations such as You Only Look Once (YOLO) v3, v4, and v5. They trained YOLOv4 and v5 algorithms on publicly available soybean dataset to recognize and discriminate the presence of weeds on the farmland. Their simulation results have shown that YOLOv5 delivered the best weed detection accuracy with a mean average precision of 96%.

The outcomes highlighted the efficacy of our approach in precise discrimination between crop-weed and detection of various weed species across diverse agricultural scenarios. The findings shows that of our proposed approach is effective for accurate detection of weed in a given field. This method is effective for controlling bermudagrass, crabgrass, and pigweed weeds in sesame fields. We evaluated our method using a dataset of 1300 images cultivated in *Kharif* season of sesame fields of China with weed infestation levels ranging from 0% to 80%. Infestation was measured as the percentage of field area covered by weeds, determined through visual assessment of each image.

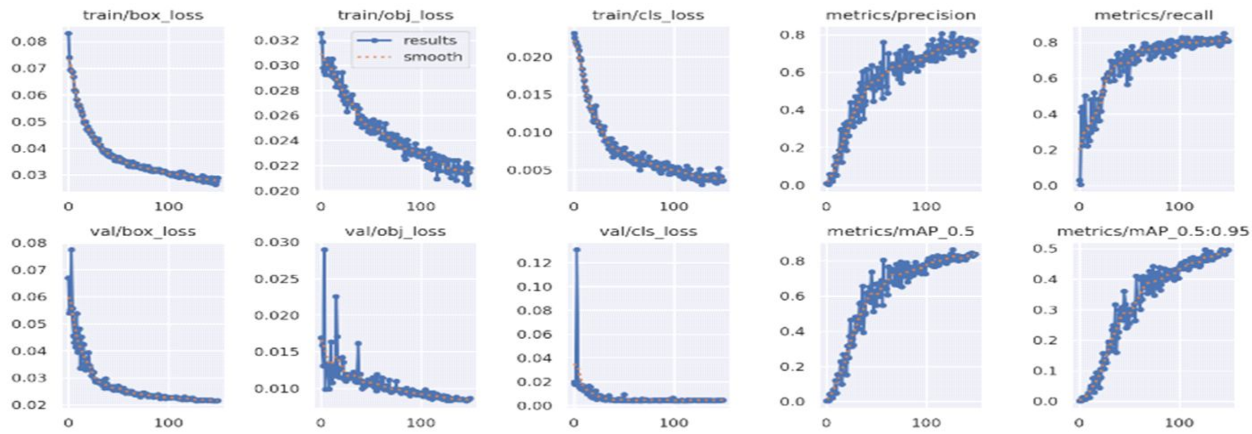

Figure 6. Graphs of YOLOv5 during fine-tuning

This demonstrates the potential for our approach to be used as a tool for farmers to manage their fields more effectively by reducing the need for manual work in weed control. Furthermore, it has potential use when integrated with other precision farming equipment, making it a cost-effective solution for farmers. With precise selective herbicide spraying, we can control the detected weeds. Overall, proposed weed-plant detection system presented a promising solution for the agricultural industry, with the potential to improve crop yields and reduce environmental impact.

Our future work will focus on classification of the weeds into different species. Future efforts will aim to integrate our system into agricultural equipment for real-time weed detection in the field. Additionally, combining our approach with other precision farming techniques has the potential to optimize crop yields and further minimize environmental impact.

REFERENCES

Ahmad A, Saraswat D, Aggarwal V, Etienne A, and Hancock B. 2021. Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. *Computers and Electronics in Agriculture* **184**: 106081.

Chen J, Wang H, Zhang H, Luo T, Wei D, Long T and Wang Z. 2022. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion, *Computers and Electronics in Agriculture* **202**: 107412.

Dhruw D, Sori AK, Tigga S and Singh A. 2023. Weed detection in soybean crop using YOLO algorithm. In Machine Intelligence Techniques for Data Analysis and Signal. pp. 777–787. Processing: *Proceedings of the 4th International Conference MISP 2022*, Volume 1. Singapore: Springer Nature Singapore.

Elhoseny, El-Beltagy, Mostafa AA and Tolba, AM 2023. Weed detection and control in precision agriculture: A review. *Sensors* **23**(1): 132.

Alotaibi SHM, Elhoseny MAA and Tolba AM. 2022. Deep learning-based weed detection in precision agriculture: A review. *Computers and Electronics in Agriculture* **189**: 106657.

Jialin Yu, Shaun M Sharpe, Arnold W, Schumann and Nathan S Boyd. 2019. Deep learning for image-based weed detection in turfgrass. ELSEVIER, *European Journal of Agronomy* **104**: 78–84.

López-Correa JM, Moreno H, Ribeiro A and Andújar D. 2022. Intelligent weed management based on object detection neural networks in tomato crops. *Agronomy* **12**: 2953. https://doi.org/10.3390/agronomy12122953

Miller JD, Singh VP (2018). The future of weed management: Challenges and opportunities. *Crop Science*.

Nath CP, Singh RG, Choudhary VK, Datta Debarati, Nandan Rajiv and Singh SS. 2024. Challenges and alternatives of herbicide-based weed management. *Agronomy* **14**(1): 126.

Nurudeen AR, Larbi Asamoah, Tanzubil Paul, Kizito Fred and Hoeschle-Zeledon I. 2024. Plant density and variety effect on yield, leaf spot disease, weed species richness and diversity of groundnut production in northern Ghana. *Weed Biology and Management*.

Pallottino F, Ceccarelli M and Piana F. 2022. Weed25: A deep learning dataset for weeds identification. *Frontiers in Plant Science* **1**: 885780.

Shanmugam S, Jeyaraj K and Ramalingam S 2021. A survey of weed detection methods based on computer vision. *ScienceDirect* **104**: 102059.

Singh AK and Gupta RK. 2022. *The Impact of Weeds on Crop Productivity: A Review. Crop Protection*.

Sportelli M, Apolo-Apolo OE, Fontanelli M, Frasconi C, Raffaelli M, Peruzzi A and Perez-Ruiz M. 2023. Evaluation of YOLO object detectors for weed detection in different turfgrass scenarios. *Applied Science* **13**: 8502.

Tushar Hrishikesh Jaware, Vinodkumar Ramesh Patil, Chittaranjan Nayak, *et al.* 2023. A novel approach for brain tissue segmentation and classification in infants' MRI images based on seeded region growing, foster corner detection theory, and sparse autoencoder. *Alexandria Engineering Journal* **76**: 289–305.

Wang CY, Liao HM, Wu YH, Chen PY, Hsieh JW, and Yeh IH. 2020. CSPNet: A New Backbone that can Enhance Learning Capability of CNN, *Conference on Computer Vision and Pattern Recognition Workshops*, Seattle, WA, USA, 1571-1580.