Network Modeling Analysis in Health Informatics and Bioinformatics (2026) 15:55
https://doi.org/10.1007/s13721-025-00726-8

ORIGINAL ARTICLE

Effective detection of weeds in sesame crop

Nawazish Ali[1] · Kashfia Sailunaz[2] · Said AbuShaar[3] · M. Kemal Özdemir[4] · Jon Rokne[5] · Reda Alhajj[4,5,6]

Received: 10 November 2025 / Revised: 5 December 2025 / Accepted: 6 December 2025
© The Author(s) 2026

Abstract
Rapid improvements in Computer vision (CV) and Deep Learning (DL) methods have enabled information extraction from images and videos to apply and interpret them for various applications. Numerous agricultural and farming applications have been using DL approaches for resolving agricultural challenges as well as optimizing farming operations. One such significant task is automated weed identification and categorization to improve crop yield. With the recent research and experiments providing information on the adverse effects of the use of chemical herbicides in farming, it is becoming evident that they are damaging the environment while wasting resources. Differentiating weeds from crops before applying herbicides can help with reducing the environmental effects. Detecting weeds in crops from images and videos is a challenging task due to the similarities between the colors, forms, and textures of weeds and crops throughout the growing period. In this paper, a deep Convolutional Neural Networks (CNN) is implemented for accurate detection of weeds in sesame crops for selective chemical herbicides spraying. Three different pre-trained deep learning algorithms like You Only Look Once (YOLO), specifically YOLOv4, YOLOv5, and Faster R-CNN are applied and the detection models achieved mean average precision (mAP) of 74%, 56%, and 52%, respectively. YOLOv4 performed the best among the three models on a limited dataset. The average IoU value outperformed other object detection-based DL models applied for similar tasks with the score of 80%, showing the efficiency of the proposed system.

Keywords Weed detection · Sesame crop · Agriculture · Computer vision · Deep learning · YOLO

1 Introduction

The detection of weeds has evolved into an essential component of precision farming, which makes use of the Internet of Things (IoT) framework. Weeds in the crops impose a threatening effect on living things, agriculture productivity, and the environment. As weeds compete with crops for resources, they are now the cause of 45% of the crop losses that occur in the agricultural industry. They reduce species diversity and disturb the harmony of ecological communities in the environment (Miller et al. 2010; Adkins and Shabbir 2014; Gharde et al. 2018). They have some severe financial effects including herbicide expenses and production losses due to weeds competing with crops for light, nutrients, water, and other critical resources. An Australian grain production system, for example, has an annual weed cost of $3.3 billion, which includes $2.6 billion in weed management expenses and $0.7 billion in yield losses (Llewellyn et al. 2016). Several traditional techniques are employed to tackle this issue and improve crop yield. One of the main techniques to remove weeds from the crop is chemical spraying and the use of herbicides. The weed control management in the conservation crop system exclusively relies on herbicides. Chemical spraying on the crop is expensive, increases air pollution, and is dangerous for living creatures. To lessen the negative impacts, weed control methods including harvest weed seed suppression (Chauhan

✉ Reda Alhajj
alhajj@ucalgary.ca

1 School of Computer Science, University of Galway, Galway H91 FYH2, Ireland

2 Department of Computer Science, Dalhousie University, Halifax, Nova Scotia, Canada

3 Middle East University, Airport highway, Amman, Jordan

4 Department of Computer Engineering, Istanbul Medipol University, Istanbul, Turkey

5 Department of Computer Science, University of Calgary, Calgary AB T2N 1N4, Canada

6 Department of Heath Informatics, University of Southern Denmark, Odense, Denmark

Published online: 17 February 2026

et al. 2012) and crop competition (Walsh and Powles 2014) are being considered to reduce the likelihood of herbicide resistance. This strain on herbicides in global agricultural systems is exacerbated by the cheap cost of the herbicides, high biochemical selectivity, simplicity of application, immediate control, and the absence of effective alternatives.

Numerous studies have shown a strong association between crop yield loss and weed competition. Lanini and Strange (1991) examined season-long competition in California lettuce crops and concluded that weed affected the production by reducing the lettuce yield up to 50%. Similarly, Hodgson (1968) showed that two Canadian thistle stem buds per square meter reduced wheat yield by 15%. Similarly, Monaco et al. (1981) reported that when the density of Johnsongrass, High Morning Glory, Common Cockle, and Crabgrass was 11 weeds/m^2, the yield of directly sowing tomatoes was reduced by 71%, 67%, 48%, and 48%, respectively. Additionally, the presence of weeds during the harvest stage can reduce the quality and value of certain crops. According to Nave and Wax (1971), harvesting soybeans before weeds dried up caused significant threshing and separating losses when the speed was raised from 1 to 2 and 3 mph. A significant increase in stuttering, lodging, and stalk losses was seen in the plots with pigweed and foxtail infestations and resulted in the greatest potential loss (34%) to the agricultural output (Oerke 2006).

To control weeds, various operations have been taken, among which chemical weed control is the most widely used method since the 1940 s. However, conventional chemical weed control sprays herbicides uniformly throughout the field, leading to the overuse of herbicides, which further leads to catastrophic environmental pollution problems. To decrease this impact, the technique of selective spraying comes in. Selective spraying plays an important role in weed control in crops. For selective spraying, the automatic detection of weeds is important. Utilizing various machine learning and deep learning techniques has been extensively used for investigating various aspects in the agricultural and farming domain. Johnson et al. (2024), conducted a survey on the computer vision techniques used for crop monitoring. Another survey was conducted in Gao and Su (2024) on weed management methods for herbaceous field crops. Chen et al. (2024) developed an approach which optimized the YOLOv8 architecture for agricultural object detection. Image analysis techniques on satellite images for a variety of discoveries and predictions within the farming domain was discussed in Rodriguez et al. (2024), whereas a demonstration on how computer vision techniques can be successfully used in the automation of various monitoring and discovery aspects in the agricultural domain was presented in Brown et al. (2024). In Pacal and Işık (2025), a framework was developed that integrated convolutional neural networks and vision transformers for precise corn leaf disease identification.

Few researchers focused on automated weed detection motivated by the major negative effects they have on crops. Hasan et al. (2024) concentrated on the detection and classification of weed species by considering object-level benchmarking based deep learning methods. An optimized DINO model was presented in Wang et al. (2025) for accurate object detection of sesame seedlings and weeds. Islam et al. (2025) considered multi-stage weed detection and classification based on WeedSwin hierarchical vision transformer with SAM-2. In Naik and Chaubey (2024), the researchers employed region-based convolution neural networks for weed detection and classification in sesame crops. Most researchers approached weed detection from image data and applied various machine learning techniques for the task. Hung et al. (2014) used a sparse autoencoder to detect three different species of weeds and achieved a precision of 94.3%, 92.9%, and 72.2%. They identified the weeds using high-resolution aerial photographs but could not localize the weeds from the images. They achieved a precision above 90% by only focusing on the classification of weeds. Cheng and Matson (2015) applied machine learning techniques like SVM, deep neural networks, and Decision Trees for the detection of weeds and rice on a publicly available dataset. They distinguished rice and weeds in the dataset without pinpointing the location of weeds and rice plants in the images. They achieved 98% classification accuracy by focusing on a binary classifier to identify weeds and rice. A fuzzy decision-making and shape descriptors for the classification of weeds type was introduced in Herrera et al. (2014). The proposed model achieved a 92% accuracy on a small dataset containing 66 images. They categorized different types of weeds but failed to recognize where the weeds resided in the pictures. Siddiqi et al. (2014) applied SVM for the classification of weeds and they achieved an accuracy of 98% on 1200 image data. As their dataset was smaller, the results lacked generalization among large amount of crop image datasets without experimentation. In a similar study, Ishak et al. (2009) achieved 93% accuracy for the classification of grass and broadleaf weeds using a single-layer perceptron on the dataset of 400 images. They used a single-layer neural network for the classification of different weeds plants and crops with high accuracy. In dos Santos Ferreira et al. (2017), the authors applied CNN for the detection of weeds in soybean plants and classified the weeds into grass and broadleaf using a convolutional neural network achieving 99% accuracy. Their experiments included SVM and other machine learning algorithms, but they did not approach localization of weeds in the images.

Various types of image features were tested with ML and DL models for weed classification and segmentation tasks.

Bakhshipour et al. (2017) used texture features from the wavelet sub-images for the segmentation of weeds. They implemented an algorithm to extract features from the sub-images using the wavelet feature and fed them to the neural network. The neural network used the extracted features and segmented the sub-divisions into weeds and crops. Only 70 images were used to classify weeds and crops, and they considered only the root mean square error to compare the performances. Like most other relevant researches, they only focused on the classification of weeds and crops without detecting their locations. Lavania and Matey (2015) proposed two methods, detection of crop rows in the images and second, distinguishing between the crops and the weeds. They used the OTSU method to convert greyscale images into binary images by calculating threshold with principal component analysis (PCA) for the identification of crops rows, and classification between crops and weeds. A sequential Support Vector Machine (SVM) was implemented in Rumpf et al. (2012) for the classification of small grain weeds by using the shape parameter from the image's dataset. They considered two types of weeds in the barley crop and achieved an accuracy of 97%. They also focused on the classification and did not consider the localization of weeds and crops for selective spraying.

Wendel and Underwood (2016) proposed a self-supervised machine learning technique for detecting the weeds in vegetables from hyperspectral grounded image data trained on a customized dataset. The classification model achieved an accuracy of 90% precision, 92% recall, and 91% recall, slightly less compared to other similar studies. A deep convolution neural network was used in Dyrmann et al. (2017) for the detection of weed locations in cereal crops. They trained the convolution neural network and validate the model performance on the 17,000 annotations of weeds from images. The implemented algorithm was able to detect only 46% of the weeds even after training their model with a comparatively larger dataset. They used GoogleLeNet for the localization of weeds in the cereals crop. Another deep learning approach was implemented for real-time weed detection using CNN and ResNet-50 on a manually generated dataset of onion and cucumber crops (Badhan et al. 2021). In Islam et al. (2021), an early weed detection system was introduced using machine learning. They used different pre-trained ML models like SVM, Random Forest (RF), and K Nearest Neighbor (KNN) for the identification of weeds from UAV images taken from Australian Chili Farms. They achieved achieved an accuracy of 96%, 94%, and 63% for SVM, RF, and KNN respectively. In Khurana and Bawa (2021), morphological scanning and textural feature analysis were applied to the images of sugar beet plants. They used a KNN classifier to classify the weeds and crops with the extracted features. Jin et al. (2021) introduced multiple techniques for the detection of weeds in vegetables. First, they trained the CenterNet model that detected only the vegetable plants and considered all other plants as weeds. Hasan et al. (2021) reviewed the existing deep learning techniques for the detection of weeds using four main steps- data gathering, data preprocessing, deep learning techniques used for the localization and classification of weeds in crops, and the evaluation metrics.

Although a good amount of research has been found on ML and DL-based weed detection from images, most studies concentrated on categorization rather than weed localization. Weed detection alone is not sufficient without their localization for selective spraying of herbicides. Precise weed identification is critical since it supplies the control system with the essential single target information for spraying, hence reducing the negative impacts on crop production and environment. This paper addressed this specific research gap in image-based weed detection problem. The major contribution of this research is weed position localization along with identifying weeds from crop images. In this paper, we employed three pre-trained CNN-based models- Faster R-CNN (Ren et al. 2017), YOLOv4 (Bochkovskiy et al. 2020; Ultralytics 2021a), and YOLOv5 (Ultralytics 2021b) to identify weeds from image data. The YOLOv4 model achieved the highest performance scores amon all three, hence was chosen for further experimentation. YOLOv4 implemented as an accurate and quick technique not only for multiple object identification but also for multiple object localization. The model was trained, validated and tested on a publicly available image dataset containing sesame plant images.

The rest of the paper is organized as follows- Sect. 2 represents the methodology of the proposed system, Sect. 3 shows the experimental results on a benchmark dataset, and finally Sect. 4 concludes the paper with potential future research directions.

2 Materials and methods

The proposed system for automated weed detection from images consists of three major components- (i) data preprocessing; (ii) ML algorithm selection and training; and (iii) evaluating chosen ML model for performance assessment. The complete methodology shown in Fig. 1 is explained in this section.

2.1 Dataset

The goal of this research is to use deep learning algorithms to identify weeds and crops in sesame crop and a publicly available image dataset (Utsav et al. 2021) was used for

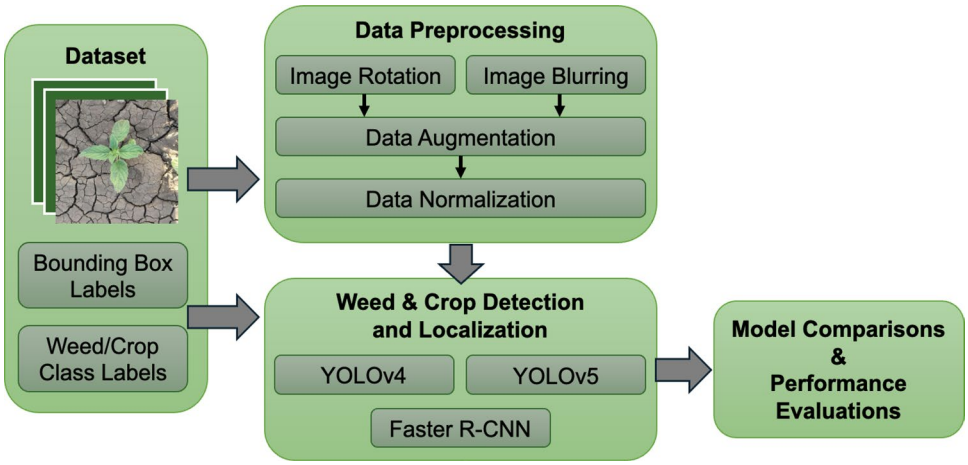


Fig. 1 Overall summary of the workflow of the complete system

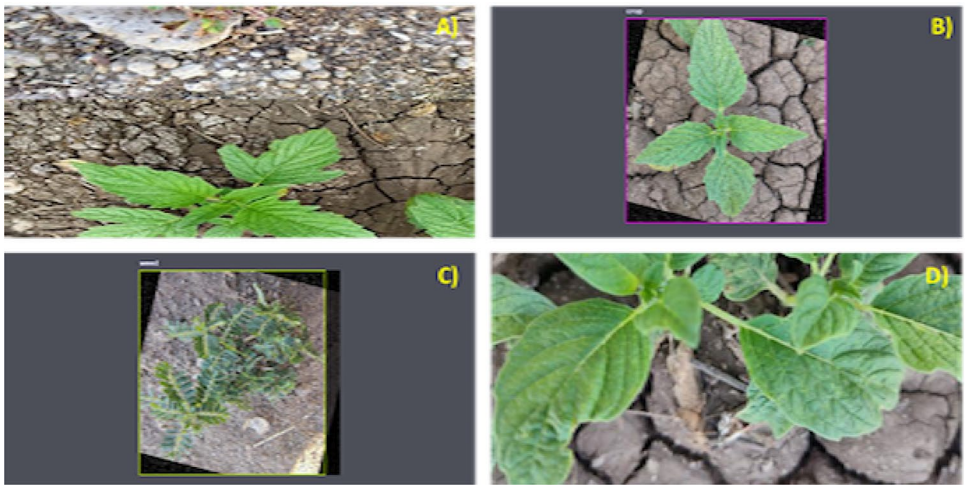


Fig. 2 A samples data image with associated preprocessing

algorithm training and validation. The collection contained 1300 images with dimensions of 512 X 512. Each image included one or more weeds and crops, hence the dataset had almost 2000 labeled objects (i.e., weed or crop) of different classes. The collection also included bounding box labels and class labels for each image with the labels in YOLO format. The dataset was divided into two classes: weeds and crops, where weed images depicted several forms of weeds and each image had at least one item from each class.

2.2 Data preprocessing

The dataset was divided into three categories- training, validation, and testing with 70:20:10 ratio as it is one of the common data division ratios among similar object identification researches on small/medium sized datasets (Raza et al. 2025). This division provided a balanced learning by maximizing the training data, as well as reduced overfitting with sufficient validation data, and finally tested the performance on unseen data using the test dataset. After the division, the training set contained 910 images with various labels, whereas the validation set and testing set had 260 images and 130 images, respectively. The testing data was used to evaluate the trained model's performance. After dividing the dataset, we employed basic data cleaning and data augmentation methods to enhance the dataset's size and add nonlinearity to the data. In the data enhancement method, images were rotated by 45 degrees and blurred by 5%. The image blurring was applied by using the Gaussian blur with a kernel size 5x5 to introduce a light blur effect on the images. Both operations used simple OpenCV functions (i.e., Rotate, and GaussianBlur). The data augmentation was applied to increase the dataset size as well as to incorporate realistic low/medium quality photos from different angles. Additionally, data normalization method was applied to enhance the algorithm's performance. During data loading, the images were converted from pascal format to dataframes and normalized, and the labels were converted into simple numerical format. Then both the converted images and labels were loaded to the system for ML model training. Fig. 2 presents examples of a sample data. Figure 2(B) and 2(C) illustrate the blurring effect and rotation angle, respectively.

2.3 Weed detection algorithms

Many applications employ CNNs for pattern and object detection and identification. In this paper, three most popular object identification CNNs named YOLOv4, YOLOv5, and Faster R-CNN were chosen for the weed detection task and to compare their performances. Faster R-CNN is a region-based convolution network for object detection that was built on on R-CNN by employing refined settings to improves the training and testing speed and to increase the performance of the network. The speed of the Fast R-CNN is 10 times faster than the previously implemented algorithms and more accurate than the previous algorithms. It is implemented using python and C++. Fast R-CNN takes the whole image as input and creates a proposal of the object. It processes the input image into convolution and max-pooling layers and creates the features map from the image. Then, it extracts the fixed-length feature vector from the feature map for the object proposal based on the region of interest (ROI). Each feature vector is fed into a series of fully connected layers, finally connected with two parallel fully connected layers. One of the layers calculates the SoftMax probability of the classes and calculates all their background classes. The second layer calculates the bounding box of the objects.

On the other hand, YOLO creates a grid from the input picture, dividing it into N×N cells, and then extracts the object from each cell in the grid. With each updated iteration, YOLO's network size, the number of layers, and the number of parameters raise for more refined identification. YOLOv5 is an expanded version of YOLOv4 that follows YOLOv4. YOLOv5's architecture is similar to that of YOLOv4. It is also divided into three sections: the backbone, the neck, and the head. The Backbone uses the input picture to extract features. In YOLOv5, CSPNets are combined with Darknet to form CSPDarknet, which serves as the backbone. It improved processing speed and accuracy over Darknet53, which was used in YOLOv3 and YOLOv4. The second component is the Neck, which builds feature pyramids that assist the YOLOv5 in scaling various sizes of objects. The neck uses the PANet (Path Aggregative Network) for various detector levels. PANet-generated conforming feature pools link the feature pyramid to multiple feature levels. The useful information is then extracted from each feature layer and sent to the subnetwork. The final detection is performed by the Head of the algorithm, which is comparable to YOLOv3 and YOLOv4. It constructs the anchor box for the feature map and returns the final vector with bounding boxes and item class probabilities.

YOLOv4 is a deep neural network model which is the extension of YOLOv3. It can detect multiple objects from images or videos. The input of the YOLOv4 is an image and it output two things. The first one is the bounding box of each object in the image and the second thing is the respective class of each object. In YOLO architectures, the backbone of the model extracts feature maps from the input images. The feature maps are then passed to the neck of the model. YOLOv4 improved the neck by adding Spatial Pyramid Pooling (SPP) and Path Aggregation Network (PANet) that enhance multi-scale feature fusion. The neck then combines and refines the features from different scales and passes them to the detection head of the model. The detection head in YOLOv4 is based on YOLOv3 that predicts bounding boxes and class probabilities at three different scales. The predicted bounding boxes are improved in YOLOv4 compared to YOLOv3. Finally, the best bounding box is selected for each detected object. YOLOv4 is faster than the previous versions as it only processes the regions in which there is a high probability of having the object. Figure 3 shows the architecture of YOLOv4.

In this study, we trained YOLOv4, YOLOv5, and Faster R-CNN using the training data and validated all models for two classes using the validation data- weeds and crops. In

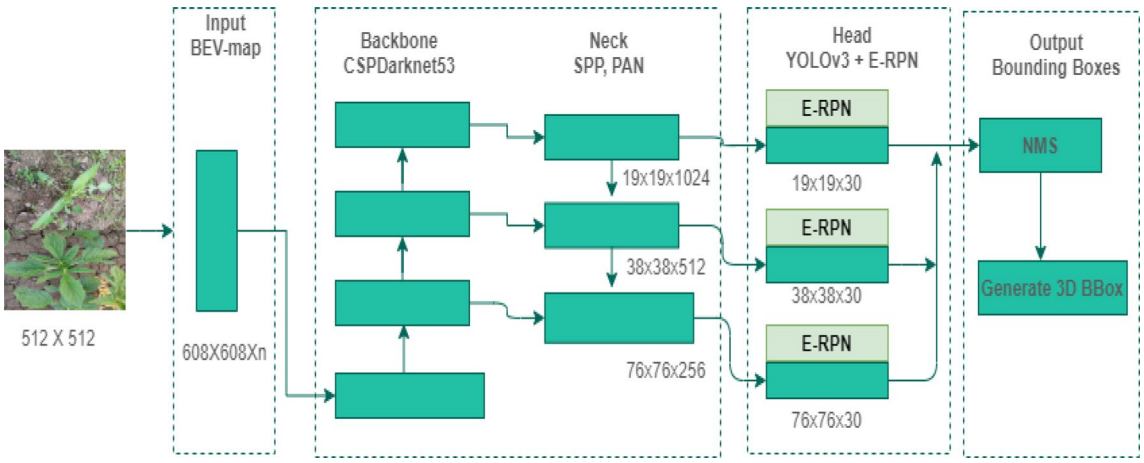


Fig. 3 Overall general architecture of YOLOv4 algorithm

Table 1 Training hyperparameters of the ML models

Models	Optimizers	Batch size	Learning rate	Iterations
YOLOv4	Adam	4	0.001	500 Epochs
YOLOv5	Adam	4	0.001	500 Epochs
Faster R-CNN	SGD	4	0.001	500 Steps

the process of fine-tuning the models, hyperparameters like learning rate, optimizer, number of epochs, and batch size were employed. The pre-trained models were loaded into memory, and the training process started by unfreezing all of the model layers. As the model performances were significantly lower in detecting the weeds for smaller epochs like 50 or 100, we tried few combinations and finally modified the hyperparameters of the models and trained them across 500 epochs. The normalized dataset also helped enhancing the model performances. The base version of this code is available on Tianxiaomo (2021).

The training parameters of the three chosen models are presented in Table 1. A learning scheduler and weight decays were also included in the models for performance enhancements. Adam optimizer was chosen for YOLOv4 and YOLOv5 training, whereas Adam and Stochastic gradient descent (SGD) optimizer were tried for Faster R-CNN with SGD providing the highest results. The models were trained with a learning rate of 0.001 after testing a range of values to find the optimized one. YOLOv4 and YOLOv5 were trained for 500 epochs, and Faster R-CNN was trained for 500 steps as all models converged by then. Although the Faster R-CNN was trained for 10,000 steps first, but as the accuracy and loss converged by 500 steps, early stopping was applied to the model to avoid overfitting. Although all three models had the same batch size, the number of iterations for the Faster R-CNN model was slightly different than the other two due to the early convergence of the model. The epochs are described to specify the number of complete passes of them model through the training data, whereas a step refers to processing one batch of data. In case of the YOLO models, they required 500 epochs to learn from the training data, whereas the Faster R-CNN converged within only 500 steps to avoid any unnecessary training.

All models were trained, validated and tested and the performances for each model in detecting weeds and crops and identifying each appropriately from the images were computed with various performance metrics. The results from all three models were compared to choose the best performing model for the task to analyze and provide possible future research directions for similar studies.

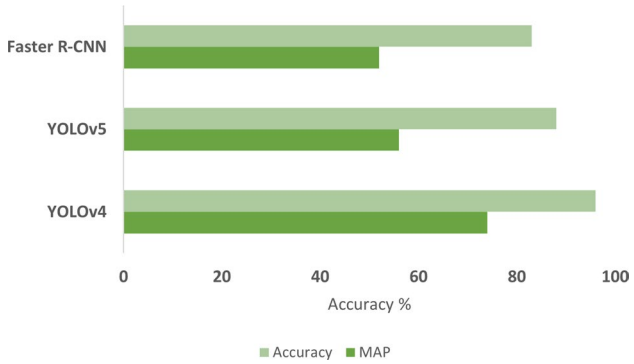


Fig. 4 Performance comparison between Faster R-CNN, YOLOv5, and YOLOv4 with the model accuracy and mAP values for weed and crop detection and classification. The light green bars represent the accuracy % whereas the dark green bars show the mAP scores

3 Results

In this paper, the performances of the trained models were measured using accuracy, Intersection over Union (IoU), confusion matrix, precision, recall, and mAP scores. As most of the existing similar studies did not incorporate localization of the crops and weeds in the images, the IoU calculation in this work increased the strength of our research. Weed and crop localization is essential for accurate detection and selective spraying. The IoU calculation on the ground truth bounding box and the detected bounding box provided an estimation on how well the implemented ML models worked for localizing the weeds.

Figure 4 shows the performance comparisons for all three models with their accuracies and mAPs scores. As obvious from the graph, YOLOv4 easily outperformed the other two models in both accuracy and mAP. The near perfect accuracy shows that YOLOv4 correctly classified almost all data images, and the high mAP represents that model provided fewer false positives and false negatives compared to the other two models. The high accuracy of the model represented the overall detection of weeds and crop showing that the YOLOv4 model was able to detect the presence of the weed and/or crop from almost all images. However, the mAP scores were lower than the accuracy scores as they were computed based on the bounding boxes of the objects (i.e., weeds and crops). Although the model accurately identified the existence of the weeds/crops, it struggled to match the exact bounding boxes and locations of the weeds/crops for some of the images lowering the mAP scores. Among the three models, YOLOv4 provided a well-balanced and accurate model for detecting, localizing, and identifying the crops and weeds from the input images.

Further analysis and performance measurements with YOLOv4 for localizations of the crops and weeds were computed with the IoU scores and the model provided an

average IoU of 80%. Figures 5 and 6 show examples of weed and crops detection and localization with the IoU scores computed from the bounding boxes (green for ground truth and red for detection). Out of 130 images, only two images had an IoU of less than 60%.

Few sample detection results of YOLOv4 are presented in Fig. 7. In the figure, the purple color bounding box represents the crop and the green color bounding box represents the weeds. In Fig. 7(a, b, c, d, e), the trained model detected one sesame plant and one weed plant, whereas Figure Fig. 7(f) shows the model detecting two sesame plants and one weed.

The results of the YOLOv4 model were significantly better in detection and localization of weeds and crops than classic ML and DL models for similar weeds/crops detection and localization tasks as mentioned in the recent researches discussed in Sect. 1. However, it failed in detection and identification of weeds and crops in some cases. Few such examples are shown in Fig. 8. In all four images, the model could not identify all of the crops and weeds from the images. As the model overlooked some plants throughout the detection process, it was unable to identify all of the plants contained within a single image. Figure 8(a) shows that the model made an accurate detection of the larger weeds; however, it was unable to recognize some of the smaller weeds. This is because the small weeds are difficult to recognize due to their size as well as the fact that the background texture of the pictures varies. Similarly, in Fig. 8(b), the model successfully detected the large size weed plant but failed to detect the small size plant. Similar cases were seen in Fig. 8(c) and 8(d). Few recurring characteristics were seen from all of these samples where the model failed to detect and identify all crops and/or weeds form the images- i) the model had trouble detecting smaller weeds/crops; ii) it could detect smaller weeds/crops when they were in close proximity compared to the ones that were scattered through the image; iii) the model struggled with weed/crop detection when their colors were more similar to the soil or background instead of bright/deep green; iv) the image quality also had an impact on the weed/crop detection as the model was not able to detect them when the crops/weeds were a bit blurry or did not have well-defined outlines; and v) sometimes the model could not detect all of the weeds/crops when there were multiple (i.e., mostly with more than two) in the image.

The model performances were further compared by calculating the precision and recall of weeds and crops with precision-recall curves. Figure 9 presents the precision-recall of crops, while Fig. 10 represents the precision-recall of weeds obtained from the classification performance of YOLOv4 for the classification of weeds and crops. The average recall for the classes were 72.75% and 75.50%



Fig. 5 A sample output detecting weed with bounding boxes for ground truth and detection with IoU 87%. Here, the green bounding box is the ground truth and the red bounding box is the detected one



Fig. 6 A sample output detecting two crops with IoU 85% and 84%. Here, the green bounding box is the ground truth and the red bounding box is the detected one

respectively. After determining the recalls, the precisions were estimated for individual classes and mAP of the two classes, and our models had a mAP of 74% for YOLOv4, 56% for YOLOv5, and 52% for Faster R-CNN. As the dataset is small, mAPs were not optimal for the three ML models. Similar to the other performance scores, YOLOv4 was chosen again due to the highest mAP. mAP for weed

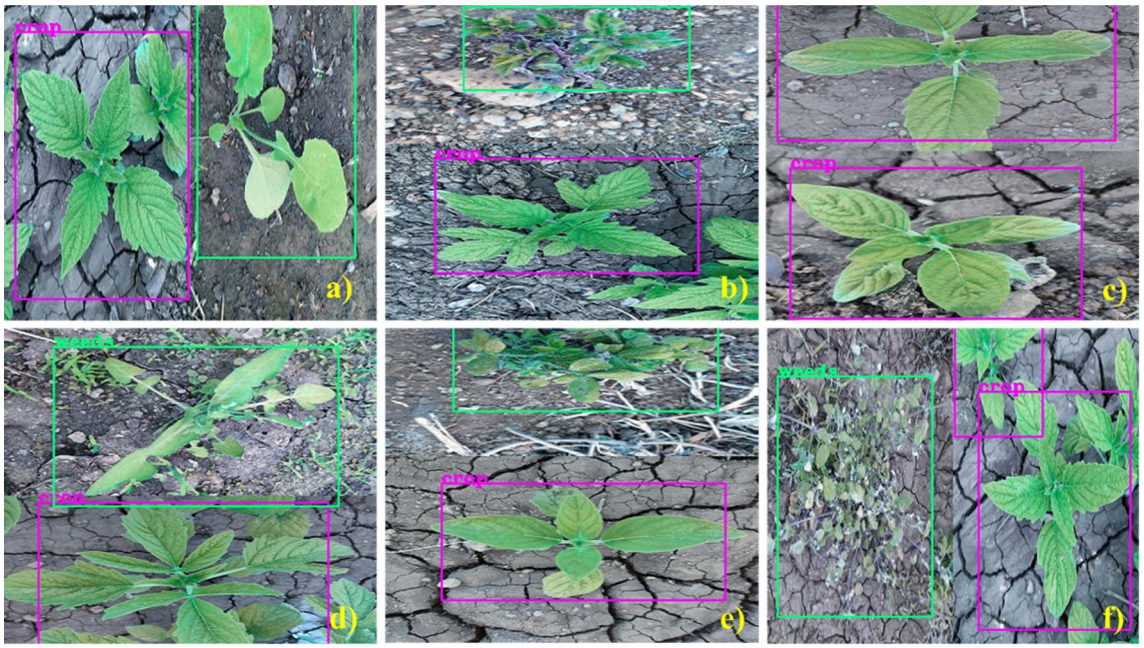


Fig. 7 Few sample detection outputs- (a,b,c,d,e) detected one weed and one crop, whereas (f) detected one weed and two crop plants. Here, the purple bounding boxes represent crops and green bounding boxes represent weeds

and crops for YOLOv4 are shown in Fig. 11 providing mAP 74.12% overall.

For the quantitative analysis of the multiple object classification of YOLOv4, the confusion matrix was generated and is presented in Fig. 12. The confusion matrix shows that the classification performance still needs improvements as a the number of false positives and false negatives are not negligible compared to the total images. One of the main reasons for this was that the classification algorithms had difficulty in distinguishing between the weeds and crop plants due the leaf sizes and colors. As expected, several characteristics of weeds and crops are similar which increased the challenges of identifying weeds and crops separately. The similarities between their shape, size, and textures of different types of plants and various types of weeds increased the challenge furthermore.

3.1 Discussion

Due to the fact that weeds have a tendency to absorb various nutrients and water from crops, and applying herbicides have adverse environmental effects, identifying weeds from crops has been an extremely important area of study within the agricultural sector. Numerous approaches to weed detection have been developed to address this issue, ranging from manual to computer vision-based systems. This research investigates using an edge-based computer vision system that uses images as data to detect weeds within a sesame crops. A variety of deep learning object detection algorithms (i.e., Faster R-CNN, YOLOv4, YOLOv5) were investigated to successfully categorize weeds present in images from the plantation. After analyzing and comparing the performances of all three models, the best results were achieved from the YOLOv4 model, hence chosen for both weed detection and localization tasks. As discussed before, most researchers focus on the classification of weeds and crops but not on the localization of weeds. For selective spraying, it is necessary to localize the weeds so that the system can automatically spray on the weeds and other crops are not affected.

Although we could not compare our YOLOv4 model accuracy scores directly to the classic ML models accuracy scores due tot he basic difference of the accuracy computations, the overall performance of our model still represented a high performing model for crops/weeds detection and localization. For example, few traditional ML models like SVM, KNN, and RF for weed detection in Australian chili farms (Islam et al. 2021) achieved very high accuracy of 96%, 94%, and 63% for different species of weeds without any weed localization. While the accuracy scores were computed differently and could not be compared, our model added the weed localization and analyzed different dimensions of the model performance. Additionally, our research considered few supplementary performance metrics to represent and analyze the results further to provide more complete and informed outcomes. Our model generated mAP and IoU for understanding the specifications of the results of both detection and localization. For instance, Jin et al. (2021) used CenterNet for weed detection in vegetable plants and achieved the highest accuracy of 85% and precision of 75%. Compared to their model, ours performed

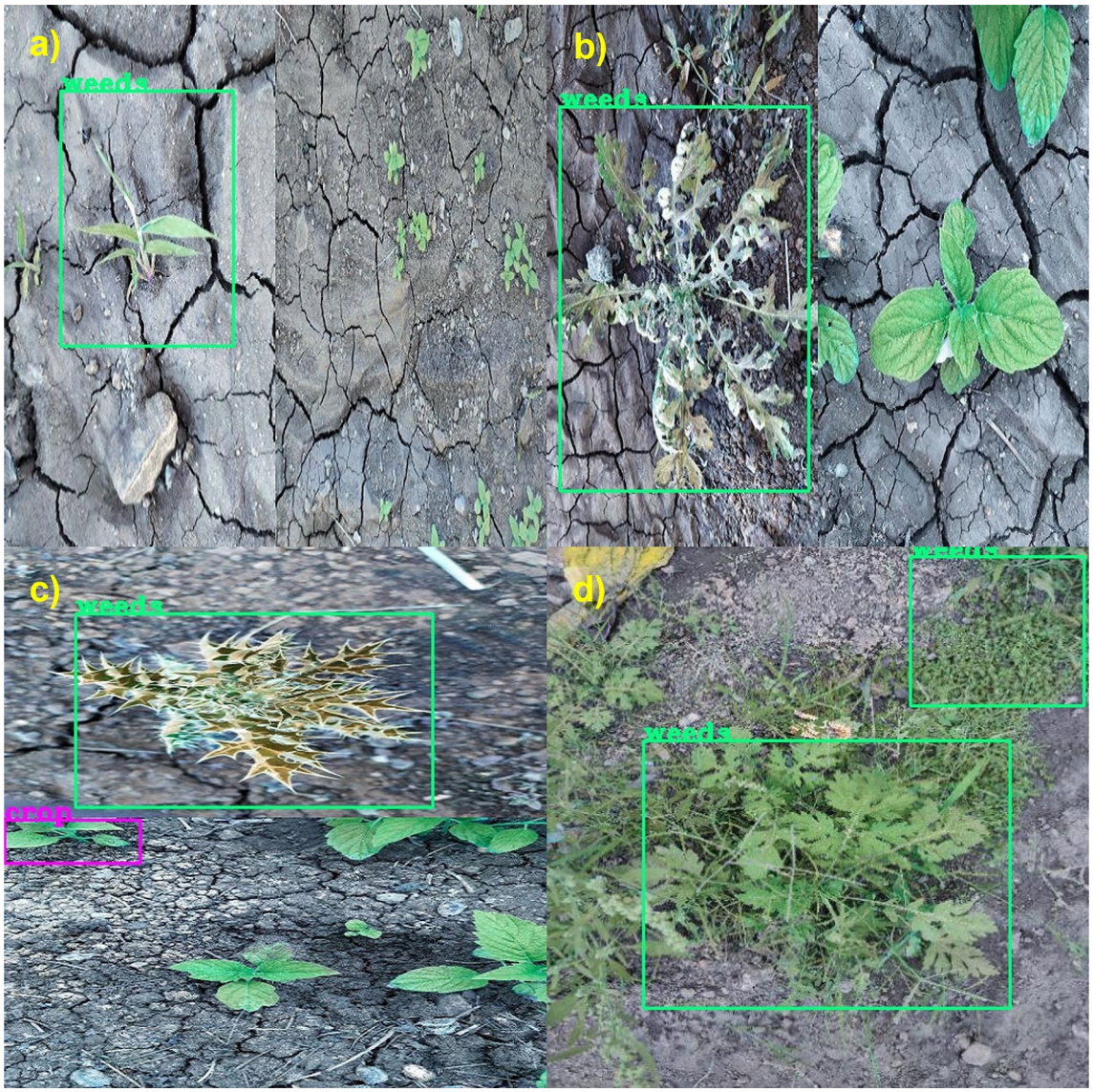


Fig. 8 Few sample outputs where the models failed to detect all crops and weeds from the images. Here, the purple bounding boxes represent crops and green bounding boxes represent weeds

better overall with 96% accuracy, 74% mAP, and an average IoU of 80%. The YOLOv4 model implemented in this paper significantly outperformed most of the similar previous researches with high accuracy, mAP, and IoU while providing a complete and detailed performance overview of the computer-vision model that enhances the reproducibility and transparency of the system.

However, the proposed model underperformed in some cases affecting the generalizability of the system. The model could not identify weeds and/or crops with smaller sizes, smaller leaves, color and/or texture similar to background, and many crops/weeds in images. The performance issues in aforementioned cases can be resolved with larger and diverse training datasets, additional preprocessing including various types of data augmentation techniques, and supplementary feature engineering for differentiating backgrounds from crops/weeds in a better way. The proposed system provides a groundwork on weed identification and localization with high accuracy and precision and additional enhancements can be applied to build on this system for better and more accurate outcomes that can be applied in real-time agricultural systems.

4 Conclusions

In this paper, three popular object detection and localization algorithms- YOLOv4, YOLOv5, and Faster R-CNN- were investigated, trained and tested for weed identification from the harvests of the sesame plant. After testing all of them,

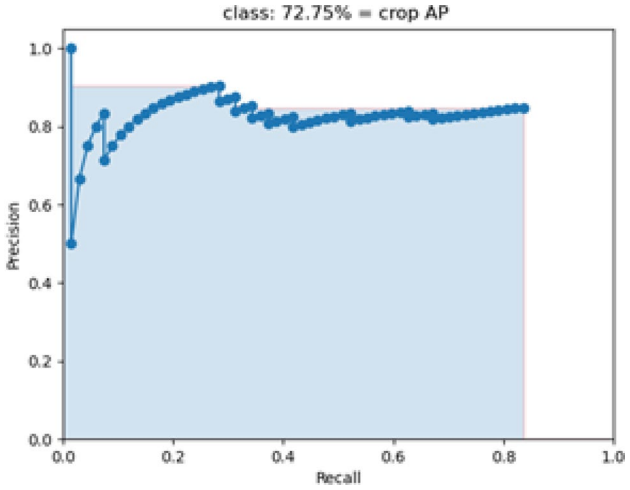


Fig. 9 YOLOv4 Precision and recall curve for crops

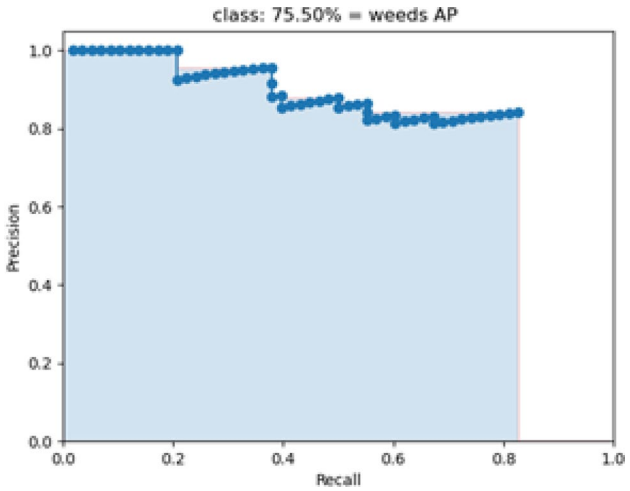


Fig. 10 YOLOv4 Precision and recall curve for weeds

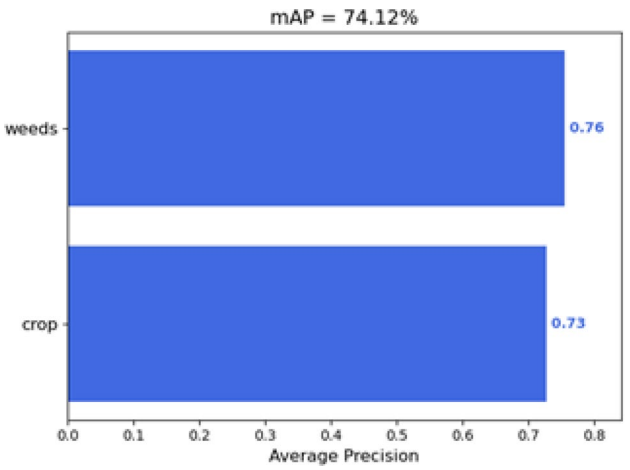


Fig. 11 YOLOv4 mAP for weeds and crops

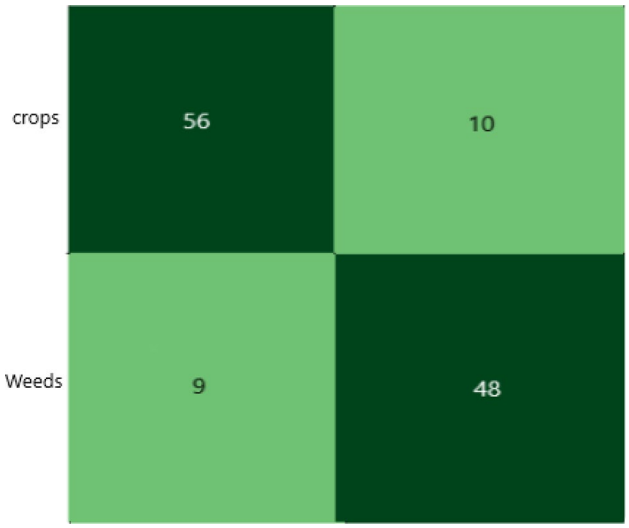


Fig. 12 Confusion matrix for YOLOv4 model performance

evaluating their performances using several metrics, and analyzing their outcomes, we chose the best model for automated weed identification and localization from image data. One of our trained algorithms (i.e., YOLOv4) has shown promising results, which are unmatched by other works in the relevant literature. As our work used the weed bounding boxes to compute mAP and IoU of the results, the overall performances from all of them (i.e., confusion matrix, accuracy, mAP, IoU) confirmed the applicability of YOLOv4 as the most appropriate model among all three. The high performance scores represent the potential implementation of such models in various agricultural farms for accurate identification and localization of weed for targeted herbicides spraying.

On the other hand, bounding box detection in the images also highlighted a few issues that need to be addressed and will be investigated in our future work. Despite achieving high performing results, additional focus is needed to enhance the performance of the weed categorization. Hence, possible future work will also include detecting and localizing all types of weeds among any type of crops. As both of these future works need larger and diverse datasets, finding and/or accumulating datasets with highest variations and data will be another challenging future research scope. Our plan to achieve such dataset includes collecting real farm data by collaborating with farms to ensure high data quality. Finally, our long-term goal is to develop a complete system on the hardware device known as Raspberry Pi that can be used in real time to capture images and immediately identify weeds for real-time applications.

Funding Open access funding provided by University of Southern Denmark

Data Availability A publicly available image dataset (Utsav et al.

2021) was used for algorithm training and validation.

Declarations

Ethical Approval This article does not contain any studies with human participants or animals performed by any of the authors.

Conflict of interest Authors declare that the research is free from any potential Conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Adkins S, Shabbir A (2014) Biology, ecology and management of the invasive parthenium weed (parthenium hysterophorus L.). Pest Manag. Sci. 70(7):1023–1029. https://doi.org/10.1002/ps.3708

Badhan S, Desai K, Dsilva M, Sonkusare R, Weakey S (2021). Real-time weed detection using machine learning and stereo-vision. Maharashtra, India, pp 1–5. https://doi.org/10.1109/I2CT51068.2021.9417989

Bakhshipour A, Jafari A, Nassiri SM, Zare D (2017) Weed segmentation using texture features extracted from wavelet sub-images. Biosys Eng 157:1–12. https://doi.org/10.1016/j.biosystemseng.2017.02.002

Bochkovskiy A, Wang CY, Liao HYM (2020) Yolov4: optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, https://doi.org/10.48550/arXiv.2004.10934

Brown T, Mann B, Ryder N (2024) Computer vision applications in agricultural automation. Artif Intell Agric 7(2):156–171

Chauhan BS, Singh RG, Mahajan G (2012) Ecology and management of weeds under conservation agriculture: a review. Crop Prot 38:57–65. https://doi.org/10.1016/j.cropro.2012.03.010

Cheng B, Matson ET (2015). A feature-based machine learning agent for automatic rice and weed discrimination, vol 9119. Springer International Publishing, Cham, pp 517–527. https://doi.org/10.1007/978-3-319-19324-3_46

Chen L, Zhang K, Liu H (2024) YOLOv8 architecture optimization for agricultural object detection. pp 89–94

dos Santos FA, Freitas DM, Da Silva GG, Pistori H, Folhes MT (2017) Weed detection in soybean crops using ConvNets. Comput Electron Agric 143:314–324. https://doi.org/10.1016/j.compag.2017.10.027

Dyrmann M, Jørgensen RN, Midtiby HS (2017) RoboWeedSupport-detection of weed locations in leaf occluded cereal crops using a fully convolutional neural network. Adv Anim Biosci 8(2):842–847. https://doi.org/10.1017/S2040470017000206

Gao WT, Su WH (2024) Weed management methods for herbaceous field crops: a review. Agronomy 14(3):486. https://doi.org/10.3390/agronomy14030486

Gharde Y, Singh PK, Dubey RP, Gupta PK (2018) Assessment of yield and economic losses in agriculture due to weeds in India. Crop Prot 107:12–18. https://doi.org/10.1016/j.cropro.2018.01.007

Hasan AM, Sohel F, Diepeveen D, Laga H, Jones MG (2021) A survey of deep learning techniques for weed detection from images. Comput Electron Agric 184:106067. https://doi.org/10.1016/j.compag.2021.106067

Hasan AM, Diepeveen D, Laga H, Jones MG, Sohel F (2024) Object-level benchmark for deep learning-based detection and classification of weed species. Crop Prot 177:106561. https://doi.org/10.1016/j.cropro.2023.106561

Herrera PJ, Dorado J, Ribeiro Á (2014) A novel approach for weed type classification based on shape descriptors and a fuzzy decision-making method. Sensors 14(8):15304–15324. https://doi.org/10.3390/s140815304

Hodgson JM (1968) The nature, ecology, and control of Canada thistle. Agricultural Research Service, US Department of Agriculture, USA

Hung C, Xu Z, Sukkarieh S (2014) Feature learning based approach for weed classification using high-resolution aerial images from a digital camera mounted on a UAV. Remote Sens 6(12):12037–12054. https://doi.org/10.3390/rs61212037

Ishak AJ, Hussain A, Mustafa MM (2009) Weed image classification using Gabor wavelet and gradient field distribution. Comput Electron Agric 66(1):53–61. https://doi.org/10.1016/j.compag.2008.12.003

Islam N, Rashid MM, Wibowo S, Xu CY, Morshed A, Wasimi SA, Moore S, Rahman SM (2021) Early weed detection using image processing and machine learning techniques in an Australian chilli farm. Agriculture 11(5):387. https://doi.org/10.3390/agriculture11050387

Islam T, Sarker TT, Ahmed KR, Rankrape CB, Gage K (2025) Weed-Swin hierarchical vision transformer with SAM-2 for multi-stage weed detection and classification. Sci Rep 15:23274. https://doi.org/10.1038/s41598-025-05092-z

Jin X, Che J, Chen Y (2021) Weed identification using deep learning and image processing in vegetable plantation. IEEE Access 9:10940–10950. https://doi.org/10.1016/10.1109/ACCESS.2021.3050296

Johnson M, Williams R, Davis S (2024) Computer vision for crop monitoring: state-of-the-art and future directions. J Agric Technol 28(2):112–128

Khurana G, Bawa NK (2021) Weed detection approach using feature extraction and KNN classification. Advances in Electromechanical Technologies. Lecture Notes in Mechanical Engineering. Springer, Singapore, pp 671–679. https://doi.org/10.1007/978-981-15-5463-6_60

Lanini WT, Strange ML (1991) Low-input management of weeds in vegetable fields. Calif Agric 45(1):11–12

Lavania S, Matey PS (2015) Novel method for weed classification in maize field using Otsu and PCA implementation. pp 534–537. https://doi.org/10.1109/CICT.2015.71

Llewellyn R, Ronning D, Clarke M, Mayfield A, Walker S, Ouzman J (2016) Impact of weeds in Australian grain production. Report for Grains Research and Development Corporation. Canberra, ACT, Australia https://www.weedsmart.org.au/app/uploads/2024/03/grdc_weeds_review_r8.pdf.pdf

Miller G, Friedel M, Adam P, Chewings V (2010) Ecological impacts of buffel grass (cenchrus ciliaris L.) invasion in central Australia-does field evidence support a fire-invasion feedback? Rangel J 32(4):353–365. https://doi.org/10.1071/RJ09076

Monaco TJ, Grayson AS, Sanders DC (1981) Influence of four weed species on the growth, yield, and quality of direct-seeded tomatoes (lycopersicon esculentum). Weed Sci 29(4):394–397. https://doi.org/10.1017/S0043174500039874

Naik NS, Chaubey HK (2024) Weed detection and classification in sesame crops using region-based convolution neural networks. Neural Comput Appl 36(30):18961–18977. https://doi.org/10.1007/s00521-024-10231-0

Nave WR, Wax LM (1971) Effect of weeds on soybean yield and harvesting efficiency. Weed Sci 19(5):533–535. https://doi.org/10.1017/S0043174500050608

Oerke EC (2006) Crop losses to pests. J Agric Sci 144(1):31–43. https://doi.org/10.1017/S0021859605005708

Pacal I, Işık G (2025) Utilizing convolutional neural networks and vision transformers for precise corn leaf disease identification. Neural Comput Appl 37(4):2479–2496. https://doi.org/10.1007/s00521-024-10769-z

Raza H, Bakr MA, Khan SD, Batool H, Ullah H, Ullah M (2025) Benchmarking YOLO models for crop growth and weed detection in cotton fields. AgriEngineering 7(11):375. https://doi.org/10.3390/agriengineering7110375

Ren S, He K, Girshick R, Sun J (2017) Faster R-CNN: towards real-time object detection with region proposal networks. IEEE Trans Pattern Anal Mach Intell 39(6):1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Rodriguez P, Martinez A, Garcia J (2024) Satellite imagery analysis for precision farming applications. Remote Sens Agric 12(4):445–462

Rumpf T, Römer C, Weis M, Sökefeld M, Gerhards R, Plümer L (2012) Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsium arvense and Galium aparine. Comput Electron Agric 80:89–96. https://doi.org/10.1016/j.compag.2011.10.018

Siddiqi MH, Lee SW, Khan AM (2014) Weed image classification using wavelet transform, stepwise linear discriminant analysis, and support vector machines for an automatic spray control system. J Inf Sci Eng, 30(4)

Tianxiaomo. pytorch-YOLOv4, (2020) Available online https://github.com/Tianxiaomo/pytorch-YOLOv4 (accessed on 20 June 2021)

Ultralytics. Ultralytics YOLOv4. Available online: https://docs.ultralytics.com/models/yolov4/ (accessed on 20 June 2021)

Ultralytics. Ultralytics YOLOv5. Available online: https://docs.ultralytics.com/models/yolov5/ (accessed on 20 June 2021)

Utsav P, Raviraj P, Rayja M (2021) Crop and weed detection data with bounding box. Available online: https://www.kaggle.com/datasets/ravirajsinh45/crop-and-weed-detection-data-with-bounding-boxes (accessed on 15 June)

Walsh MJ, Powles SB (2014) High seed retention at maturity of annual weeds infesting crop fields highlight the potential for harvest weed seed control. Weed Technol 28(3):486–493. https://doi.org/10.1614/WT-D-13-00183.1

Wang Y, Xu SF, Ye ZY, Cheng KH (2025) Optimized DINO model for accurate object detection of sesame seedlings and weeds. Sci Rep 15:11854. https://doi.org/10.1038/s41598-025-96826-6

Wendel A, Underwood J (2016) Self-supervised weed detection in vegetable crops using ground based hyperspectral imaging. pp 5128–5135. https://doi.org/10.1109/ICRA.2016.7487717

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.