

Research Article

# Crop-Weed Segmentation and Classification Using YOLOv8 Approach for Smart Farming


Sandip Sonawane<sup>1,\*</sup> , Nitin N. Patil<sup>2</sup> 

<sup>1</sup> Department of Computer Engineering, R. C. Patel Institute of Technology, Shirpur, Kavayitri Bahinabai Chaudhari North Maharashtra University, Jalgaon, 425001, India

<sup>2</sup> Shri Vile Parle Kelavani Mandal's College of Engineering, Shirpur, Dr. Babasaheb Ambedkar Technological University, Lonere, 402103, India

\*Corresponding Author: Sandip Sonawane, E-mail: sandipsonawane2006@gmail.com

Article Info	Abstract
Article History	Accurately segmenting crop and weed images in agricultural fields is crucial for precision farming and effective weed management. This study introduces a new method that leverages the YOLOv8 object detection model for precise crop and weed segmentation in challenging agricultural scenes. Our approach involves preprocessing agricultural images to enhance feature representation, followed by YOLOv8 for initial crop and weed detection. Thorough experiments using standard datasets comprising 2630 images demonstrate the effectiveness of our proposed method concerning precision, recall, mean average precision (mAP), and F1 score compared to existing techniques. These findings contribute to advancing crop-weed segmentation techniques, offering practical solutions for efficient weed management and precision agriculture. Our proposed approach outperforms state-of-the-art methods found in the literature. Our methodology presents a promising framework for automated crop-weed segmentation with applications in crop monitoring, yield estimation, and weed control strategies, supporting sustainable agricultural practices.
Received Jul 31, 2024	
Revised Sep 13, 2024	
Accepted Sep 23, 2024	
<b>Keywords</b>	
Weeds	
YOLO	
Segmentation	
Object Detection	
Precision Agriculture	

 **Copyright:** © 2024 Sandip Sonawane and Nitin N. Patil. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY 4.0) license.

## 1. Introduction

In modern agriculture, the accurate identification and segmentation of crop and weed species from images play a pivotal role in optimising farming practices [1]. Differentiating between these entities is essential for effective crop management, weed control, and ensuring high agricultural productivity [2, 3]. Traditional manual segmentation methods are time-consuming and susceptible to human error, which underscores the importance of developing the need for automated solutions to provide efficient and accurate segmentation of crops and weed instances within agricultural imagery [4].

Recent advancements in deep learning (DL) and computer vision have spurred substantial advancements in object detection and segmentation. Among these, You Only Look Once (YOLOV8) has emerged as a powerful framework for object detection, renowned for its potential to handle real-time data while maintaining high accuracy [5]. While YOLOV8 excels in detecting objects within images, its potential for

precise segmentation of complex scenes, such as agricultural landscapes comprising crops and weeds, remains relatively unexplored [6].

This work aims to rectify this deficiency by proposing an efficient approach for crop-weed image segmentation, leveraging the strengths of YOLOV8. By integrating YOLOV8's object detection capabilities with tailored segmentation techniques, we seek to develop a methodology that accurately delineates crop and weed instances within agricultural images. Through this approach, we aim to address the challenges associated with traditional segmentation methods, including occlusions, varying illumination conditions, and multiple overlapping instances [7].

The significance of automated crop-weed segmentation extends beyond academic research; it holds profound implications for practical applications in precision agriculture and weed management [8]. Accurate segmentation enables farmers to precisely monitor crop growth, assess weed infestation levels, and implement targeted interventions, thereby minimising herbicide usage, reducing environmental impact, and optimising resource allocation [9].

This paper presents our proposed methodology for crop-weed image segmentation based on the YOLOV8 approach, along with comprehensive experimental evaluations to validate its effectiveness. We believe that our research contributes to advancing the field of automated crop management and precision farming, providing practical solutions for ecological agricultural applications in an era of increasing demand for food security and environmental conservation.

The key contributions of this paper are:

- To design and develop a robust and accurate weed segmentation model using the YOLO algorithms.
- To evaluate the performance of the proposed model in segmenting different crop-weed species commonly found in agricultural environments.
- To provide a foundation for integrating YOLO-based crop-weed classification and segmentation systems into precision agriculture practices.
- To compare the performance of proposed approaches with similar existing approaches.

## 2. Literature Survey

Weed-crop segmentation is an important computer vision task for precision agriculture [10]. It tries to automatically identify between crop plants and weeds in digital photos collected from various sources, including ground-based sensors and drones. Accurate segmentation enables targeted weed management strategies, resulting in improved crop yield, less pesticide use, and a lower environmental effect [11]. Early methods of crop-weed segmentation included conventional image processing techniques such as thresholding and colour segmentation.

Numerous research has presented novel methods for classifying crops and weeds using different image analysis techniques. To train a neural network to identify pixels as either weed or crop, Bakhshipour,

et al. [12] developed a technique that extracts 52 different texture properties from photos using the Principal Component Analysis (PCA) method to identify 14 differentiating features. Similarly, Perez, et al. [13] effectively classified crop kinds and weeds using colour information and shape analysis to distinguish between vegetation and backdrop. Using form data, Jafari, et al. [14] used RGB colour components to identify sugar beetroot plants and differentiate them from common weed varieties. Zheng, et al. [15] employed a Support Vector Machine (SVM) classifier to distinguish maize crops and their weeds by focusing on colour variables and extracting nine important vectors. Lin, et al. [16] used a Decision Tree classifier to categorise weed and maize crops by combining form and texture features. Hamuda, et al. [17] created an algorithm that distinguishes between cauliflower crop patches in photos and weeds by combining colour features with morphological approaches. Together, this research demonstrates various advanced strategies for automated machine learning and image processing-based weed and crop identification.

The above-discussed conventional image processing methods have several benefits, such as requiring less training time and little data for training [18, 19]. Furthermore, the models mentioned above do not require Graphics Processing Units (GPU) with much power [20]. However, there are a few restrictions attached to them. First, the fine intricacy of visual data may not be adequately taken by existing methodologies, which are frequently dependent on deliberately produced elements like colour, texture, and shape [21]. Second, these techniques are unreliable and can be easily affected by variations in the appearance of weed, image noise, and illumination changes [22]. Finally, individuals may find it difficult to adjust to various weed species or plant growth phases or to generalise successfully to new, unseen images [23].

Some researchers have used deep learning algorithms to investigate novel methods for crop-weed segmentation. You, et al. [24] created a DL model by incorporating DropBlock and hybrid-dilated convolution into the backbone network. By expanding the receptive area, the hybrid-dilated convolution improves the network's contextual awareness, while DropBlock helps with weight regularisation by removing contiguous image patches. Sodjinou, et al. [25] merged the K-means clustering technique with a DL-based U-Net segmentation approach to efficiently segment crops and weeds. To improve performance, Osorio, et al. [26] added the normalised difference vegetation index (NDVI) as a background subtractor to SVM with Histograms of Oriented Gradients (HOG), YOLOV3, and Mask-RCNN for identifying weed in a lettuce field. Kim and Park [27] presented a convolutional neural network (CNN)-based multi-task semantic segmentation approach that can identify weeds and crops at the same time. To categorise weeds and crops in RGB images, Fawakherji, et al. [28] presented a two-network sequence that uses a segmentation approach with an encoder-decoder to classify every pixel into either the mask or soil. For identifying weeds in a sugar beet farm, Nasiri, et al. [29] used a U-Net segmentation model that included ResNet50 architecture for extracting features. Khan and his team [29] created a cascade framework called CED-Net for crop and weed segmentation to obtain coarse-to-fine segmentation results. These results are then merged to achieve refined

segmentation outputs. These works demonstrate the adaptability and efficiency of deep learning techniques in tackling precision agriculture's weed-crop segmentation problems.

When using artificial images created from actual plant examples and GANs, Khan, et al. [30] discovered that the artificial images outperformed real datasets using Mask R-CNN. U-Net++ outperformed the conventional U-Net in weed picture segmentation, as shown by Sapkota, et al. [31], who achieved greater accuracy and IoU metrics, with a special emphasis on small weed detection. YOLOv8 outperformed Mask R-CNN regarding orchard segmentation, showing faster inference times and improved recall and precision [32]. RDS\_Unet, a specialised semantic segmentation network designed for maize seedling field analysis, was introduced by Cui, et al. [33] and significantly improved IoU, precision, and recall metrics over U-net, an established approach.

These strategies took advantage of the unique colour qualities of crops and weeds. However, their efficiency was restricted by factors such as fluctuating light conditions, overlapping foliage, and immature weed stages with negligible colour distinction [34].

The introduction of DL, specifically CNNs, has transformed weed-crop segmentation. CNNs are excellent at extracting complicated features from visual data, making them ideal for this task [35]. Pioneering efforts used Fully Convolutional Networks (FCNs) to carry out segmentation, correctly categorising pixels in the input images as crops or weed. The U-Net architecture increased segmentation accuracy by including skip connections, which preserve spatial information necessary for exact boundary delineation [36].

The study compared various model architectures for crop-weed segmentation and classification [37], including lightweight CNNs and RNNs. Many researchers have employed machine-learning models not only for agricultural tasks but also in the medical field [38]. While lightweight CNNs are effective for classification [39], they lack accuracy and efficiency in detection and segmentation tasks. RNNs are less applicable for static image-based problems. The YOLOv8n-Seg model was chosen due to its task-specific efficiency, streamlining object detection and segmentation, and real-time detection capabilities. This model, particularly its lightweight "n" version, balances speed and accuracy, making it a practical choice for real-time, pixel-level segmentation.

### 3. Materials and Methods

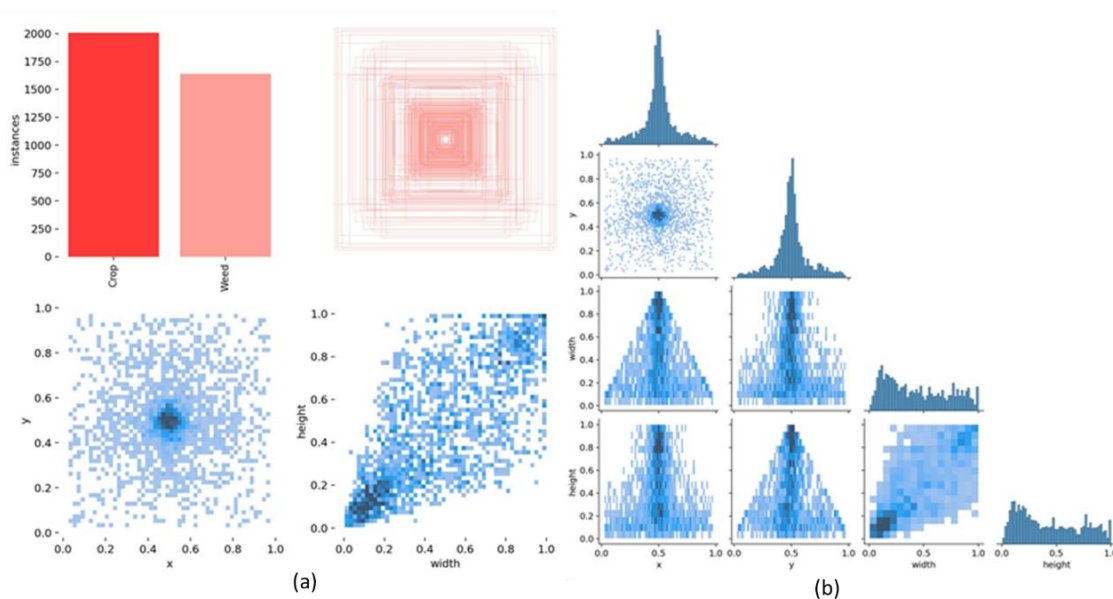
The proposed work used a dataset available on the Roboflow platform [40]. It contains 2630 images of crops and weeds. The input images in the dataset were separated into train, validation, and test subsets. Crop and weed samples in the dataset are visualised in Figure 1.

Figure 2 (a) represents the distribution of labels, where the y-axis is labelled "Instances" and specifies the number of instances. The "crop" bar is taller than the "weed" bar, suggesting that there are more instances of crops than weeds in the dataset. Overall, the figure presents data on crops and weeds, possibly

related to their spatial distribution (as suggested by the density plots) and the number of instances observed (as shown in the bar chart). The figure shows some distribution of measurements related to the crops and weeds.



**Figure 1.** Input images of weeds and crops in the dataset



**Figure 2.** Correlogram of instances in dataset (a) Distribution of labels (b) Correlation of labels

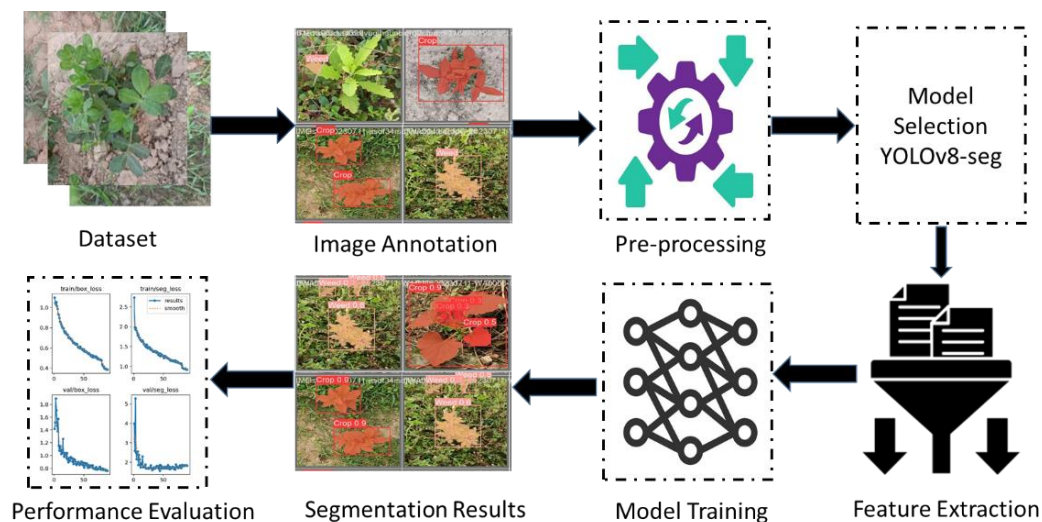
Figure 2 (b) indicates a labelled correlogram, a graphical representation of the correlation between different variables. In object detection with YOLOv8, this correlogram likely represents the relationships between the normalised coordinates and bounding box dimensions predicted by the trained model. The correlogram includes scatter plots and histograms. Scatter plots show the relationship between two variables. Histograms along the diagonal, display the distribution of a single variable. The correlogram facilitates an understanding of the object detection model behaviour and can help diagnose issues or biases in the detection patterns, such as a tendency to detect objects in certain parts of the image or of certain sizes more frequently [41].



### 3.1. Proposed Methodology

This research adopts a YOLOv8-Seg-based image segmentation pipeline shown in Figure 3. Initially, a dataset of representative images is collected and meticulously annotated with corresponding segmentation masks. Data preprocessing steps like resizing and normalisation are applied to enhance training efficiency. In resizing, we scaled all the images to the same size and used normalisation to scale pixel values to the [0,1] range, ensuring consistent input for effective training. Furthermore, random flipping, rotation, and scaling were used to replicate various orientations and sizes, enhancing the model's robustness to real-world scenarios [42]. A YOLOv8-Seg model, chosen based on factors like hardware limitations and dataset complexity, is then trained on the prepared dataset [43, 44]. The dataset is divided into subsets for training and testing. The training dataset was employed to instruct the model on segmenting objects depending on the given masks. Finally, the performance of the proposed approach is evaluated on the testing images with metrics like IoU and mAP to assess its generalisation ability.

Using sophisticated backbone and neck architecture, the YOLOv8 improves feature extraction and object detection performance. This model emphasises maintaining an optimal balance between processing speed and detection accuracy. Notably, its Anchor-Free Detection feature makes the detection process more efficient and accurate. Additionally, applying the mosaic data augmentation method expands the size and diversity of the training dataset, thereby enhancing model generalisation.



**Figure 3.** Workflow for training a YOLOv8-Seg model for image segmentation tasks

The YOLOv8n-Seg model offers multiple aspects that enhance its interpretability. One key feature is its ability to perform both box detection and mask segmentation, allowing for image visualisation and analysis of crop and weed boundaries. This dual capability clearly explains how the model distinguishes between different objects. Additionally, the model assigns confidence scores to each identified object, which helps assess the reliability of its predictions and provides insight into how certain or uncertain the model is in classifying crops or weeds. Furthermore, the architecture of YOLOv8n-Seg is designed for efficiency

and interpretability. Its simple, lightweight structure allows for easier understanding than more complex models while offering robust end-to-end detection and segmentation. This combination of visual output, confidence scoring, and straightforward architecture enhances the model's transparency and reliability.

Google Colab was utilised to train the YOLOv8n-Seg model, with the dataset images and label files prepared on the Roboflow platform and then imported into the program using API code in YOLOv8 PyTorch format. The training process was executed on a Windows 10 machine using Ultralytics YOLOv8 version 1.29 with Python 3.10.12 and Torch 2.2.1 with CUDA support on a Tesla T4 GPU. This arrangement provides enough computational resources to handle the enormous dataset and ensure efficient training. The system requirements also featured the Tesla T4 GPU with 16 GB of VRAM for enhanced deep-learning applications. The number of studies presented in the literature study have employed object detection and segmentation models like improved YOLOv5-Seg and YOLOv8-Seg for crop-weed classification, segmentation, and detection, often utilising standard and optimised hyperparameters outlined in Table 1 during model training. Therefore, these standard hyperparameters were similarly applied in our model training process. The hyperparameters provided in Table 1 were chosen through experiments rather than relying primarily on previous research. We originally tried several configurations of learning rates, batch sizes, and optimiser types to determine the ideal settings for optimising the model's performance while avoiding overfitting. The final settings, such as the starting learning rate of 0.0001 and batch size of 32, were chosen to strike a balance between training efficiency and model correctness. Additionally, the dropout rate of 0.25 and weight decay of 0.0005 were adjusted to minimise overfitting during training.

**Table 1.** Configuration of hyperparameters in YOLOv8 algorithm

Parameters	Configuration
Epochs	100
Input Image size	640x640
Optimizer	Auto (AdamW)
Batch size	32
Learning rate (lr0)	0.0001
Learning rate (lrf)	0.01
Dropout	0.25
Weight decay	0.0005
Seed	42
Batch Size	32
Momentum	0.9

### 3.2. Model Evaluation Indicators

The research employed precision (P), recall (R), mAP, and F1 score as criteria to assess accuracy. P denotes the proportion of the predicted algorithm's detected area to the actual detected area, whereas R

specifies the ratio of correctly predicted classes out of all required classes. The mAP assesses sample accuracy based on anticipated boxes exceeding half of the actual bounding boxes. An increase in mAP values indicates improved prediction accuracy.

In this manuscript, TP denotes the count of true positive samples, FP indicates the score of false positive images, N signifies the total quantity of samples, and Q represents the number of detected sesame crops. The average precision for the  $i$ -th class is denoted as  $AP_i$ . The Matthews Correlation Coefficient (MCC) assesses the quality of binary classification. It ranges from -1 to 1, with 1 indicating a flawless forecast, -1 indicating complete disagreement, and 0 suggesting no association [45]. All these metrics were computed from the confusion matrix, and the formulas for their calculation are given below.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

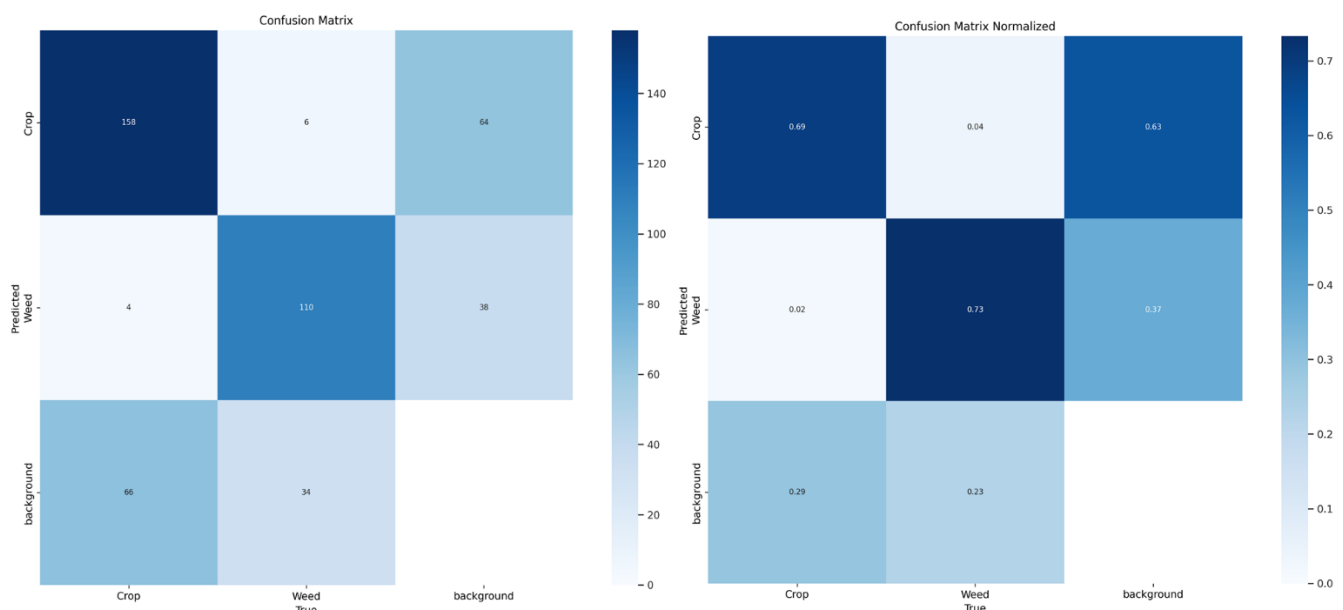
$$mAP = \frac{\sum AP_i}{n} \quad (3)$$

$$\text{IoU} = \frac{TP}{FP + TP + FN} \quad (4)$$

$$\text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (5)$$

## 4. Experimental Results

### 4.1. Confusion Matrix



**Figure 4.** Confusion Matrix illustrates the classification performance of the YOLOv8n-Seg model



Here, the confusion matrix in Figure 4 provides a more comprehensive view of the model's performance across different classes. Based on the provided confusion matrix, we calculated the precision, recall, MCC, etc.

Tables 2 and 3 provided a thorough performance evaluation of the efficiency of the proposed YOLOv8 model in box detection tasks and mask segmentation tasks, specifically for identifying boxes and masks, respectively. The model performs reasonably well, with precision, recall, and mAP scores indicating a good balance between correctly identifying objects and minimising false positives. For both box detection and mask segmentation, the 'weed' classes generally exhibit superior performance in identifying weed instances compared to other classes ('Crop' and 'All').

From Table 2, our findings suggest promising levels of accuracy in detection. The "Crop" and "Weed" classes attained a precision value of 79.3% and 84.6%, recall rates of 74.9% and 80%, and mAP@0.5 results of 82.6% and 86%, respectively. Furthermore, the overall performance, with a mAP@0.5 of 84.3%, underscores the efficacy of the YOLOv8 segmentation algorithm in facilitating accurate and efficient box detection in agricultural settings. Regarding detection, YOLOv8s-seg achieved precision, recall, and mAP at different levels of IoU (0.5 and 0.5:0.95) of 0.846, 0.8, and (0.86, 0.606), respectively, for the "weed" class.

**Table 2.** Box Detection Performance of YOLOv8 Segmentation Model

Box Detection						
Class	Image	Instances	P	R	mAP@0.5	mAP@0.5-0.95
All	220	378	0.819	0.775	0.843	0.58
Crop	220	228	0.793	0.749	0.826	0.555
Weed	220	150	0.846	0.8	0.86	0.606

Our analysis in Table 3 indicates strong detection performance for both crop and weed classifications. The "crop" and "weed" classes accomplished a precision score of 80.4% and 86.2%, a recall rate of 75.8% and 80.7%, and mAP@0.5 of 83.1% and 86.5%, respectively. Notably, the comprehensive evaluation yields an overall mAP@0.5 of 84.8%, reflecting the algorithm's robustness and accuracy in mask segmentation tasks. In terms of segmentation performance, YOLOv8s-seg demonstrated precision, recall, and mAP results for the "weed" class at various IoU thresholds (0.5 and 0.5:0.95) of 86.2%, 80.7%, and (86.5%, 55.5%), respectively.

**Table 3.** Mask segmentation performance of the YOLOv8 segmentation model

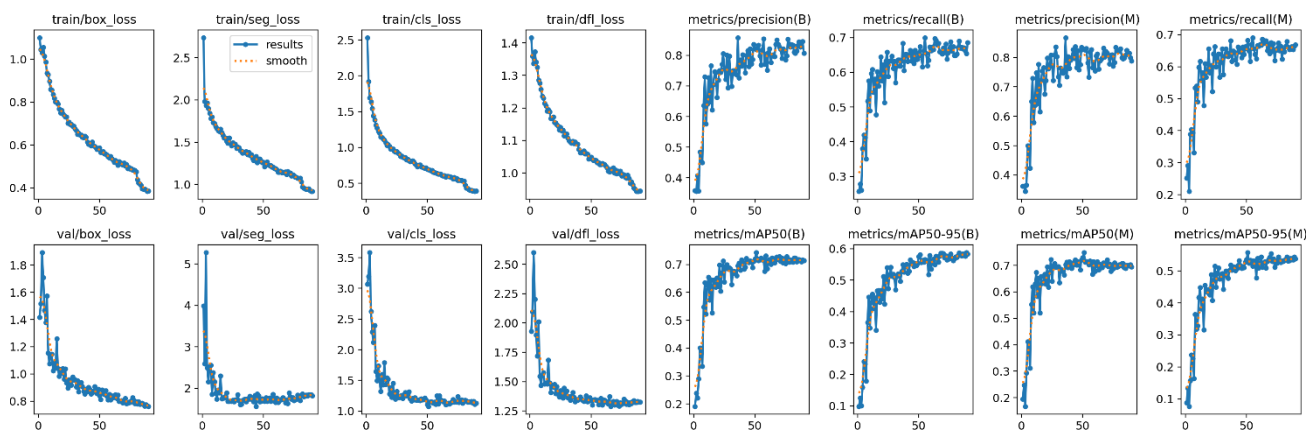
Mask Segmentation						
Class	Image	Instances	P	R	mAP@0.5	mAP@0.5-0.95
All	220	378	0.833	0.782	0.848	0.557
Crop	220	228	0.804	0.758	0.831	0.558
Weed	220	150	0.862	0.807	0.865	0.555

## 4.2. Training Curves

Figure 5 depicts a series of performance graphs associated with the training process of a YOLO v8 model, presumably for object detection and instance segmentation tasks. Each graph visualises a distinct performance metric across training iterations. The horizontal axis (x-axis) tracks the number of training epochs, whereas the vertical y-axis denotes the measured value. Blue markers denote the actual measurements, and the orange dashed line serves as a smoothed representation, facilitating trend visualisation by mitigating noise.

The loss graph examines the accuracy of predicted bounding rectangles and assesses the capability of the trained model to segment objects within the images. As illustrated in the figure, the model achieved less value for loss and maximum values for precision, recall, and mAP's for box detection and mask segmentation tasks across trained and validation datasets. A lower loss value indicates improved accuracy in box detection and more effective segmentation in the case of mask segmentation. The commonly utilised metric in the realm of object detection is mAP. The higher value of mAP represents better performance.

In conclusion, the observed decrease in loss values and increased precision, recall, and mAP metrics suggest a progressive improvement in model performance as training progresses.



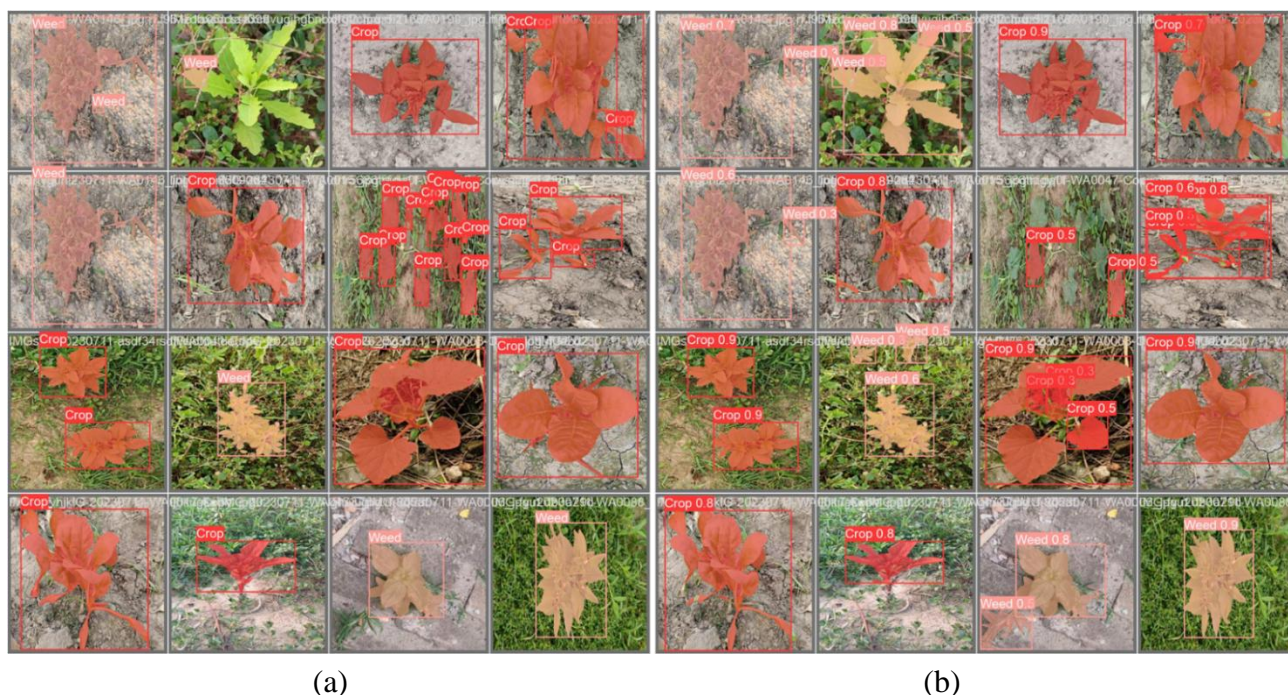
**Figure 5.** Performance metric curves YOLOv8l-Seg model for box detection and mask segmentation task

## 4.3. Segmentation Results

Figure 6 represents the output of the YOLOv8 model by identifying and classifying objects as "weed" or "crop." A bounding box surrounds each object that the model has identified. The label inside the box indicates the class of the object ("Weed" or "Crop"), and the number represents the confidence score of the detection, varying from 0 to 1, with 1 being the highest confidence.

In Figure 6(a), we see a variety of crops and their weeds with bounding boxes around them. In Figure 6(b), the confidence scores are visible next to the labels. For example, one instance is identified as "crop" with 0.9 confidence threshold, indicating a maximum of confidence in the prediction. Another instance is identified as "weed" with a 0.5 confidence threshold, which demonstrates a reasonable confidence level.

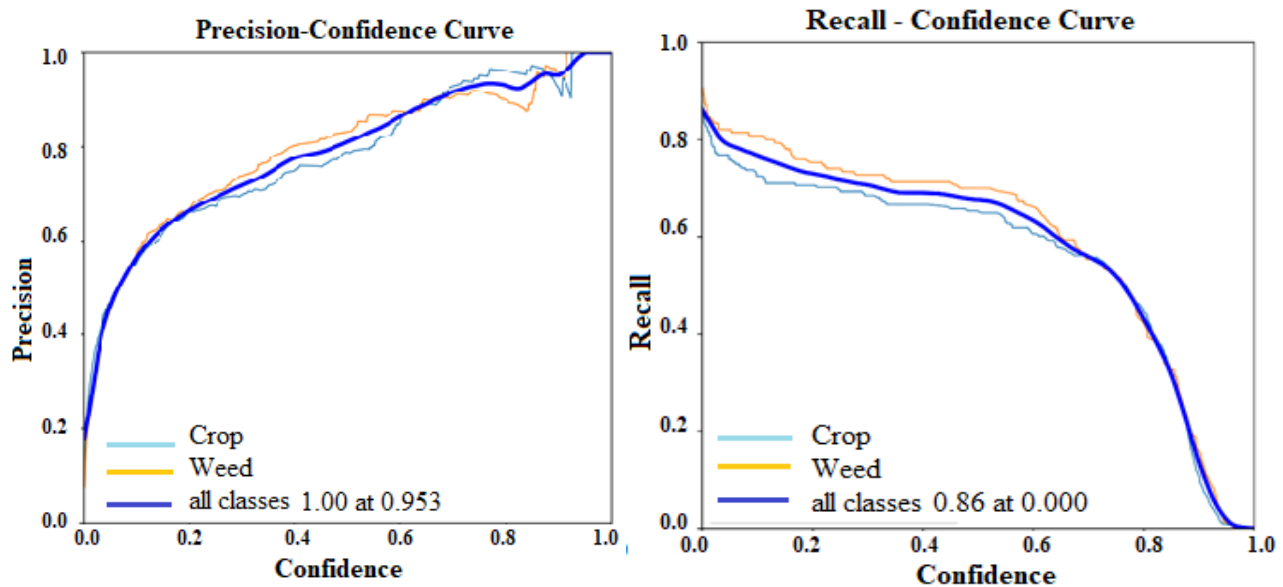
This study's detection results highlight correct classifications and misclassifications produced by the proposed model. This provides visual examples of the YOLOv8n-Seg model accurately segmenting and classifying crops and weeds, demonstrating its capabilities in detecting objects of varying sizes and placements. Additionally, we present examples of misclassifications, illustrating the model's limitations, such as confusion between overlapping crops and weeds or cases with poor lighting conditions. These qualitative results provide a more comprehensive understanding of the model's performance, helping to assess its robustness and areas for improvement.



**Figure 6.** (a) Image annotation of crops and their weeds (b) YOLOv8 predictions with a confidence score

#### 4.4. Box Detection Curves

The performance of the YOLOv8 model on a box detection task is visualised in Figure 7. The precision-confidence curve in Figure 7(a) illustrates the precision levels achieved by the model (how many detections were correct) on the y-axis against the confidence score, which ranges from 0 to 1 (the model's certainty in its prediction) across the x-axis for different classes. The display is a pair of curves, orange for "crop" and blue for "weed". These curves show the precision of the model for detecting crops and weeds at various confidence scores. The blue line marked as "all classes 1.00 at 0.953" indicates that when assessing all classes collectively, the model attains a 100% precision rate when the confidence score is 0.953. It indicates very high precision at this particular confidence level. The model performs better when the curve is higher throughout. The graph demonstrates that the model achieves excellent precision in detecting both "crop" and "weed" across various confidence levels, achieving a flawless precision score when all classes are combined at a high confidence score. It demonstrates the effective performance of the YOLOv8 model in distinguishing between crops and weeds in the given dataset.

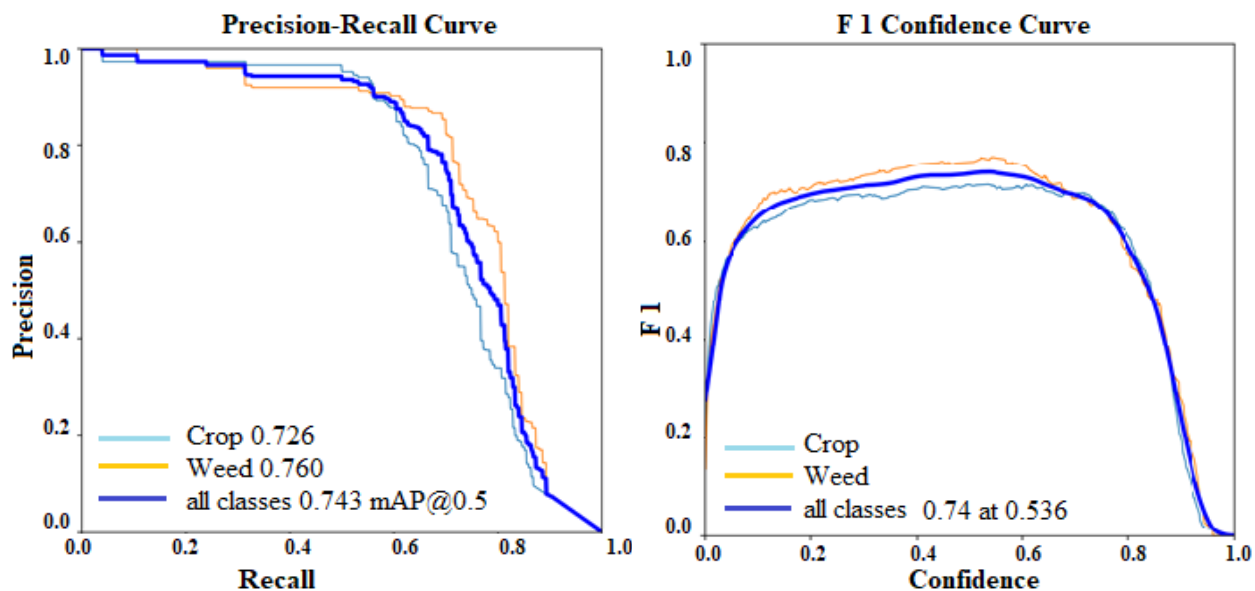


**Figure 7.** Assessment of the YOLOv8 model's efficacy in detecting objects within bounding boxes.  
(a) Precision - Confidence Curve (b) Recall - Confidence curve

The Recall-Confidence graph in Figure 7(b) assesses the model's performance regarding its capacity to correctly identify objects (recall) at various confidence threshold levels. This graph analyses the trade-off between recall and confidence for predictions on different classes of models. Here, the confidence threshold increases across the x-axis, and recall is drawn across the y-axis. The blue curve labelled "all classes 0.86 at 0.000" indicates the overall performance across all classes at a specific confidence threshold, which in this case is very low (0.000), suggesting that the recall is 0.86 when almost all detections are considered, regardless of confidence. The curves for "crop" and "weed" are quite close, suggesting that the model's performance is relatively similar for both classes. Both curves start with high recall at low confidence thresholds, which is typical as the model will predict many positives (including both true and false positives) when the threshold is low.

A Precision-Recall graph in Figure 8 (a) visually represents the association for a model for diverse thresholds between precision (y-axis) and recall (x-axis). This visualisation provides a thorough overview of how the model performs at different confidence levels. Ideally, the curve should trend towards the top-left corner, indicating a strong balance between precision and recall. The blue line in the graph is used for "crop" and the orange line for "weed." These lines illustrate the association between recall and precision for detecting crops and weeds. The legend indicates that the average precision (AP) for "crop" is 0.726 and for "weed" is 0.760. These values summarise the precision-recall curve and indicate the model's overall performance across the entire range of possible classification confidence levels. The thick blue line at the bottom of the graph depicts the mAP with a 0.743 score at 0.5 IoU threshold of 0.5 combined for all classes. The mAP of 0.743 designates good performance of the model in detecting objects when considering a moderate IoU threshold with a 0.5 value.

Overall, the graph illustrates that the proposed model demonstrates a comparatively elevated level of precision in recognising both "crop" and "weed" detections across most recall levels, with slightly better performance for "weed" detection.



**Figure 8.** Assessment of the YOLOv8 model's efficacy in detecting objects within bounding boxes. (a) Precision-Recall Curve (b) F1 - Confidence Curve

The F1-Confidence curve in Figure 8(b) illustrates the model's F1 score, with the y-axis representing this metric across the confidence threshold (certainty of the model in its prediction) on the x-axis for different classes. It represents the harmonious relations between precision and recall and offers a way to look at both metrics simultaneously. In an ideal scenario, the F1 curve would be a horizontal line at 1.0, indicating that the model achieves precision and recall perfectly. The higher the curve is throughout, the higher the model's efficacy. The orange and blue curves depict the model's F1 score for detecting crops and weeds, respectively, at various recall confidence thresholds. The model achieves a 0.74 F1 score with a 0.536 confidence value. This clarifies that the F1 score of 0.74 represents the total performance of all model classes.

Overall, the graph indicates that the model demonstrates a strong F1 score for both "crop" and "weed" detections across a range of confidence levels, with the best overall performance for all classes combined occurring at a confidence threshold of approximately 0.536.

#### 4.5. Mask Segmentation Curves

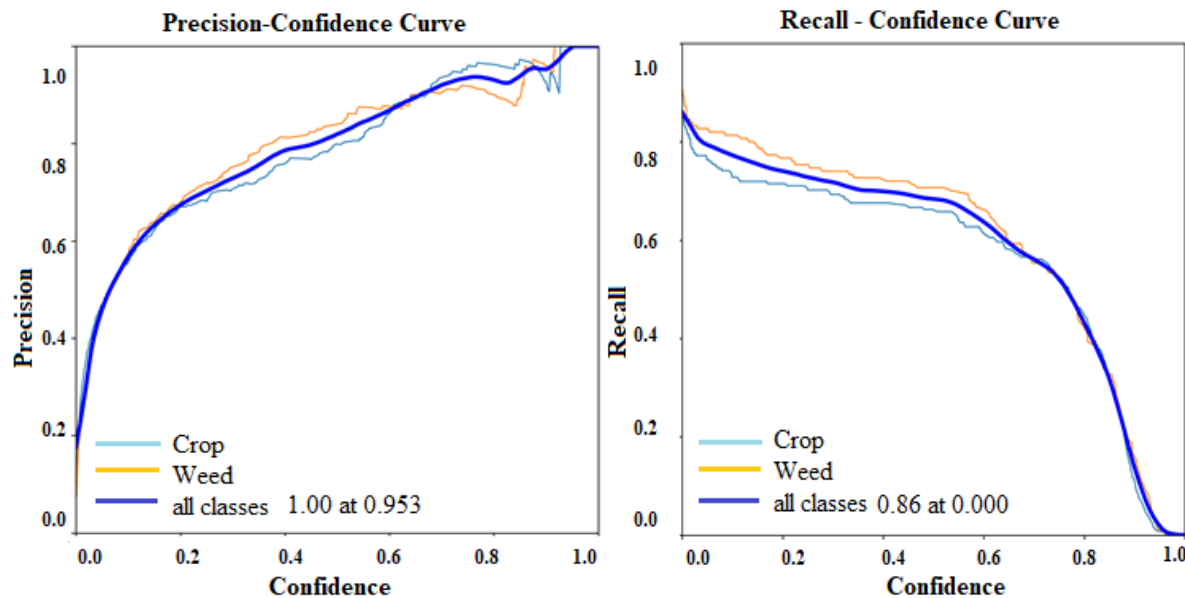
The evaluation of the YOLOv8 model's performance on a mask segmentation task is visualised in Figure 7.

##### 4.5.1. Precision-Confidence

The confidence threshold, varying from 0 to 1, is depicted along the x-axis in Figure 9(a). The level at which the model determines whether an object is present or not is known as the confidence threshold.



The y-axis represents precision, which also ranges from 0.0 to 1.0. Precision assesses the ability of a model to recognise true positives correctly. It is calculated by dividing the number of predictions the model classified as positive by the number of correct positive predictions. The orange and light blue curves indicate the precision of weed and crop class, respectively. The blue curve represents the precision of all classes combined at various confidence thresholds. The point marked on this curve with the label "all classes 1.00 at 0.953" indicates that the model correctly identifies all positive cases at a confidence level of 0.953.



**Figure 9.** Assessing the performance of the YOLOv8 model on the mask segmentation task (a) Precision-Confidence Curve (b) Recall-Confidence curve

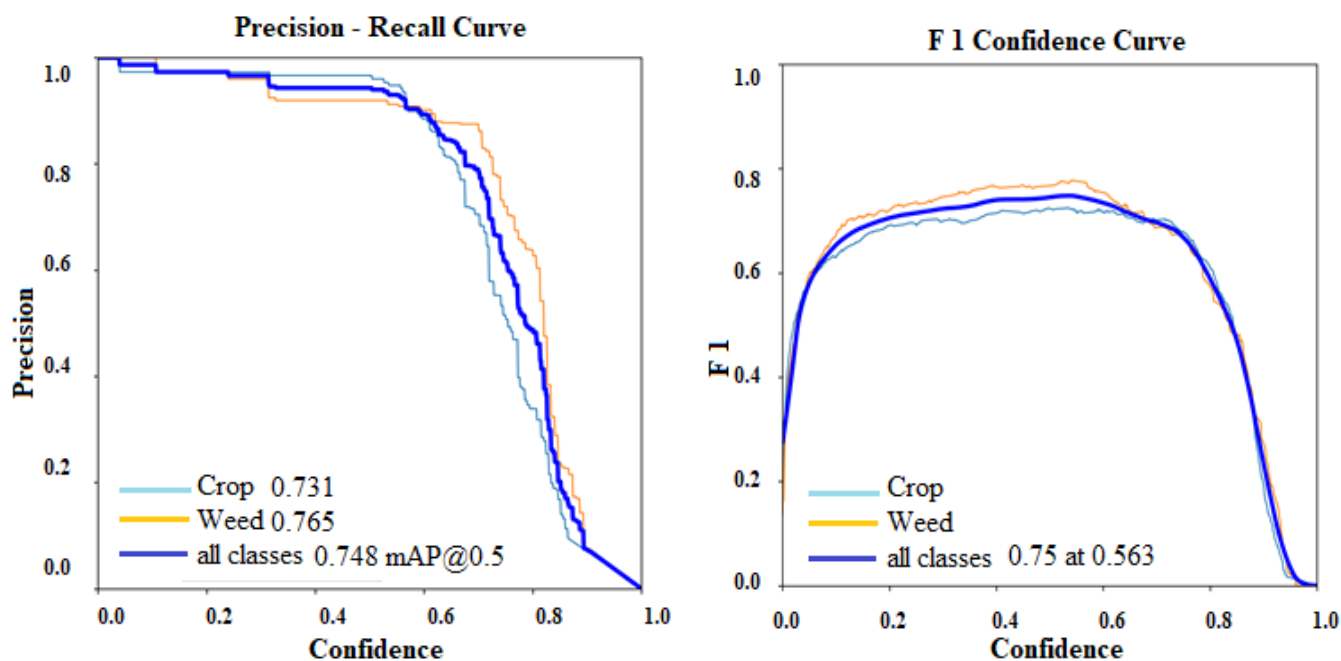
#### 4.5.2. Recall Confidence

The relationship between the model's recall and the confidence threshold used for predictions is seen in this graph in Figure 9(b). The x-axis in the curve denotes confidence score, and the y-axis represents recall. The graph includes two separate curves. The blue curve corresponds to detections of the "crop" class, while the orange curve represents detections of the "weed" class. For every class, the thick blue line shows the mean performance, indicating a specific value of 0.86 at a confidence limit of zero. This demonstrates a large recall rate, suggesting the model's capability to identify relevant portions of objects across all categories at this threshold.

#### 4.5.3. Precision-Recall Confidence

The graph in Figure 10(a) shows two curves, each representing the precision-recall relationship for different classes detected by the model. The vertical axis of the curve indicates precision, whereas the horizontal axis indicates recall. Precision indicates how accurate the predictions are when the model claims an object belongs to a certain class. Recall evaluates the ability of the model to recognise every pertinent instance of a specific class. The blue curve with an Average Precision (AP) score of 0.731 represents the

"crop" class, whereas the orange curve with an AP score of 0.765 represents the "weed" class. Overall, the graph indicates the YOLOv8 model demonstrates effective discrimination between the "crop" and "weed" categories, with slightly better performance on the "Weed" class based on the AP scores. The mAP@0.5 value of 0.748 demonstrates robust overall performance concerning the designated IoU threshold.



**Figure 10.** Assessing the performance of the YOLOv8 model on the mask segmentation task (a) Precision-Confidence Curve (b) Recall-Confidence curve

#### 4.5.4. F1 Score

The graph in Figure 10(b) illustrates the confidence threshold across the x-axis and recall across the y-axis. The blue line in the curve denotes the recall-confidence curve, which aggregates data from all classes and is labelled "all classes 0.86 at a 0.000," suggesting that at a 0.0 confidence score, the recall for all classes is 0.86. The orange and light blue lines represent the recall-confidence curves for the "weed" and "crop" classes. The lines are very close throughout the graph, suggesting that the model's performance on these two classes is comparable regarding recall across various confidence levels.

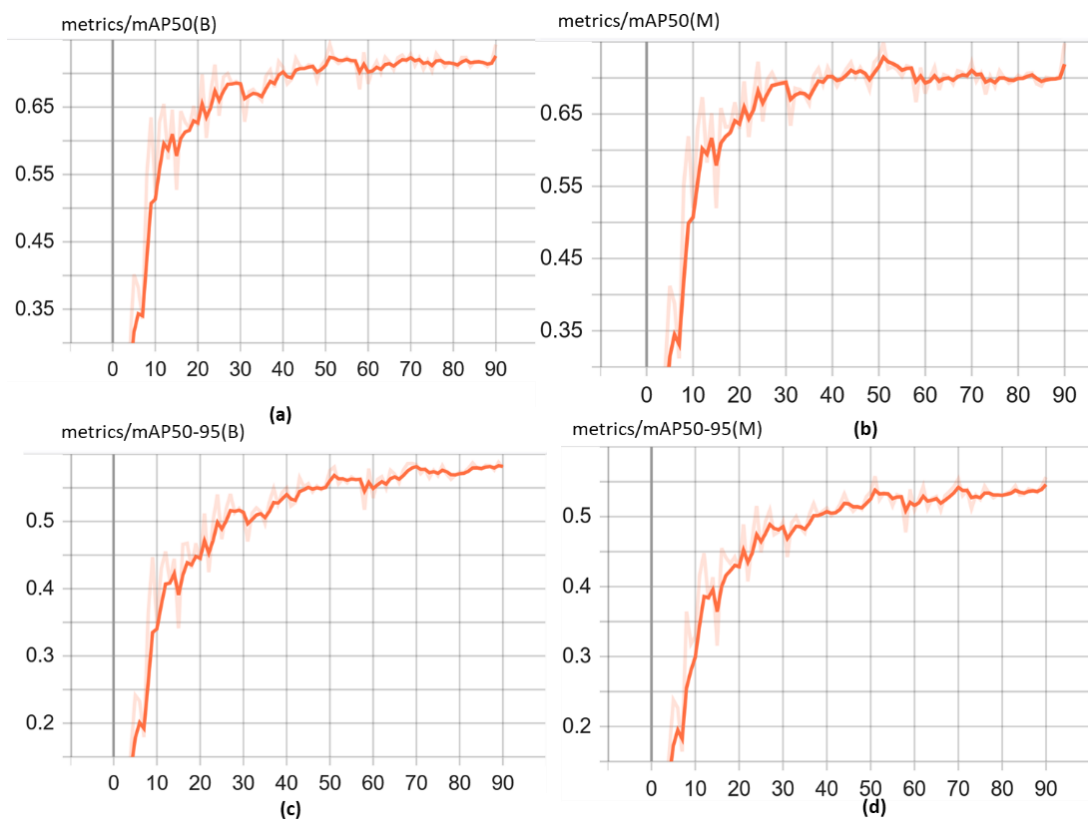
Each of the three curves originates from the upper-left corner of the chart, illustrating high recall at low confidence thresholds. With the increase in confidence threshold, there is a typical decline in recall across the board. This outcome is anticipated because the model becomes more confident about its predictions and more selective, which can lead to fewer true positives being detected, thus lowering the recall.

#### 4.5.5. Visualisation using TensorBoard mAP50 and mAP50-95 (Box and Mask)

TensorBoard is a visual aid for TensorFlow, allowing you to monitor various metrics while training YOLOv8. These charts in Figure 8 illustrate a particular model's performance outcomes. The graphs plot



different metrics over the number of epochs (or iterations) during training. The Figure 11(a) and Figure 11(b) graphs show the metric "mAP50" for the bounding box (B) predictions and mask (M) predictions, respectively. Figures 11(c) and 11(d) graph show the metric "mAP50-95" for the bounding box (B) predictions and mask (M) predictions, respectively. Across all the graphs, the horizontal axis (X-axis) tracks training progress by showing the number of training iterations completed. The vertical Y-axis shows the performance metric value being measured. The lines illustrate how these metrics change over training, and the shaded areas depict the variation or uncertainty around those values. The rising metric values across the graphs suggest that the model's effectiveness is generally increasing over time. The fluctuations and the shaded areas suggest some variance in the performance, which is typical during training due to factors like adjustments of Learning rate (LR), data augmentation, and the randomness in the model training.

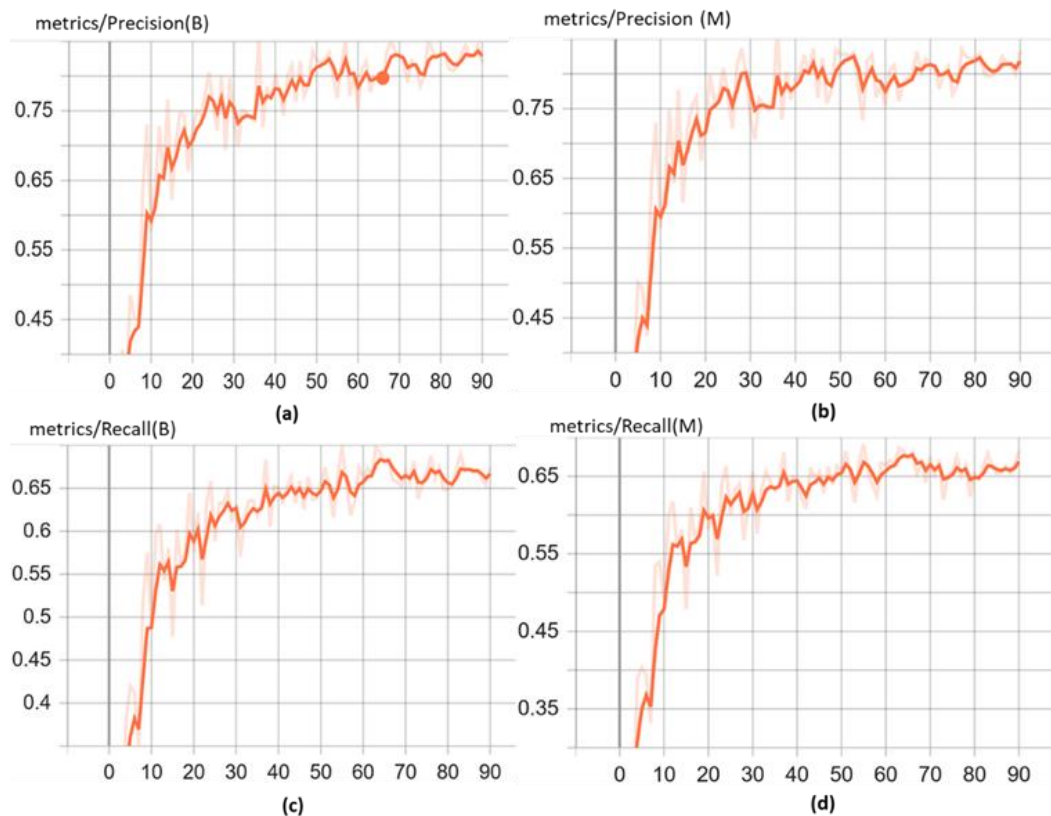


**Figure 11.** Visualisation of mAP metric using TensorBoard

#### 4.5.6. Precision and Recall (Box and Mask)

Graphs combining recall and precision visually represent two crucial metrics used to evaluate how well an object detection or segmentation model performs. Figure 12 displays each metric for both bounding box predictions (B) and mask predictions (M), relevant for tasks involving object detection and instance segmentation. The graph labelled "metrics/precision(B)" in Figure 9(a) shows the precision of the bounding box predictions over several epochs. The model is more accurate in its positive predictions as precision increases, indicating a reduction in false positives. The graph labelled "metrics/precision(M)" in Figure

12(b) shows the precision of the mask predictions. This is similar to the bounding box precision but specifically for the segmentation masks that the model predicts. In Figure 12(c), the graph labelled "metrics/recall (B)" shows the recall of the bounding box predictions. The model is more accurate in its positive predictions as precision increases, indicating a reduction in false positives. In Figure 12(d), the graph labelled "metrics/recall (M)" shows the recall for the mask predictions. This measures how well the model is at detecting the actual objects and correctly segmenting them.



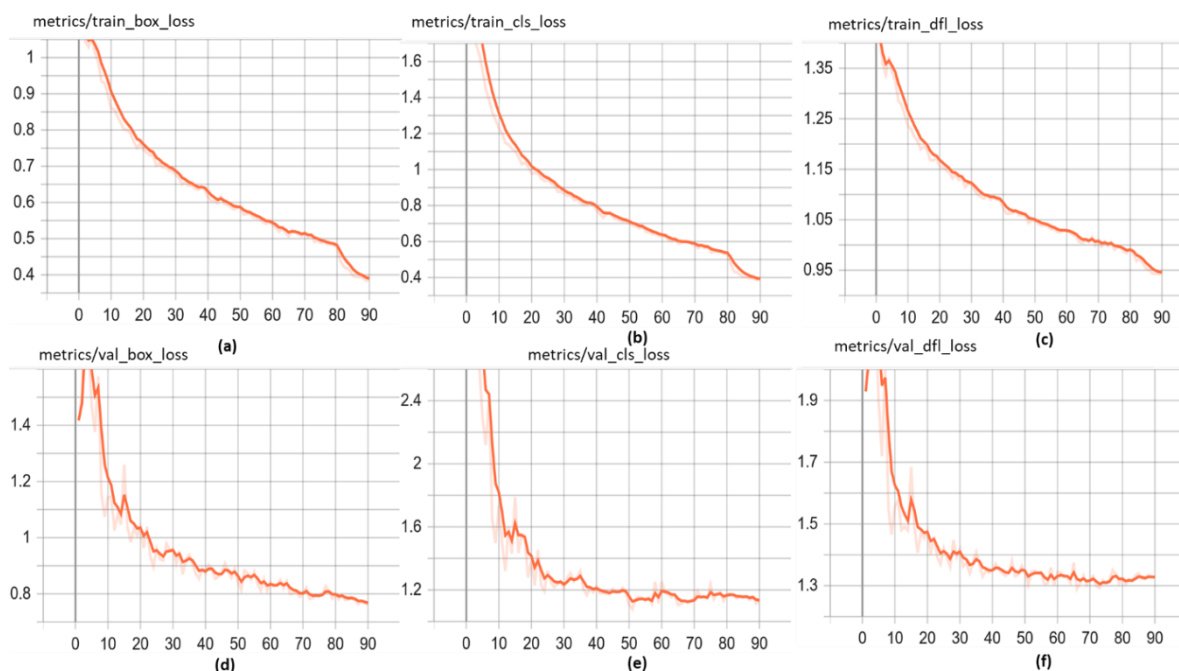
**Figure 12.** Visualisation of precision and recall metrics using Tensorboard

The graphs show how the model's performance changes over training. Each horizontal step represents one training cycle (epoch) completed on the training data. The vertical axis shows the value of the specific performance measure being tracked. The solid line represents the average performance, and the shaded area indicates how much the performance varied during training. These graphs show that precision and recall are improving over time for both bounding box and mask predictions, with performance metrics representing improvements in the model's effectiveness during training.

#### 4.5.7. Loss (Box and Mask)

In Figure 13, each graph represents a different loss metric over training epochs. It shows both datasets' box loss, classification loss, and domain-specific loss (train and validation). The model's prediction accuracy is increasing as the loss decreases. This suggests that the model is getting better at pinpointing object

locations. All six graphs consistently depict a decline in loss. This suggests the model is effectively learning and enhancing its performance using training and validation sets of data.



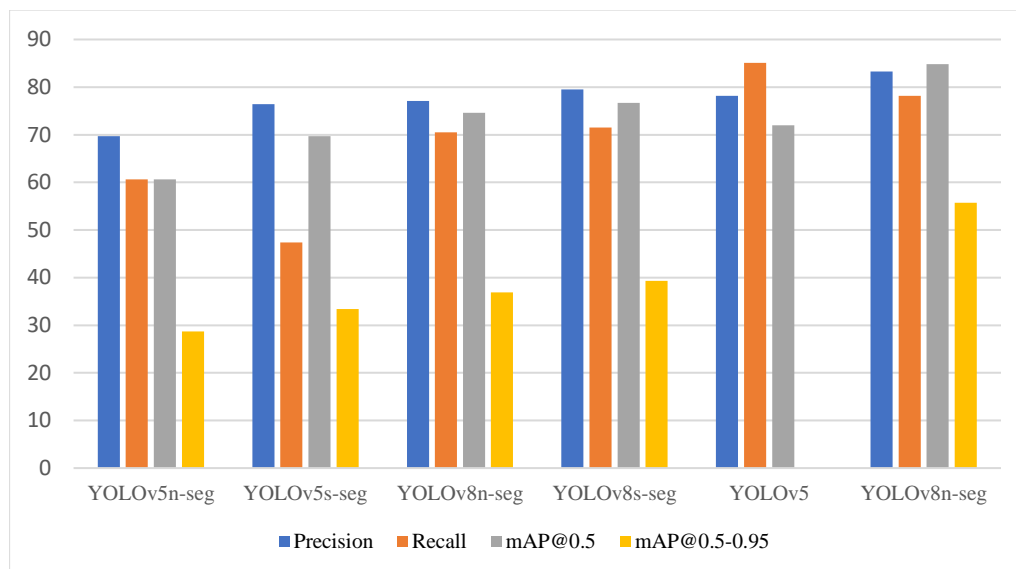
**Figure 13.** Visualisation of loss metric using Tensorboard

#### 4.6. Comparison of the Proposed Algorithm with Existing Algorithms

In this study, an evaluation of different YOLO model variants was conducted for the image segmentation, as shown in Table 4. The models assessed include YOLOv5n-seg, YOLOv5s-seg, YOLOv8n-seg, and YOLOv8s-seg, each differing in architecture complexity and model size. Performance metrics such as precision, recall, F1 score, mAP@0.5, and mAP@0.5-0.95 were employed to compare the models. Remarkably, YOLOv8n-seg and YOLOv8s-seg demonstrated superior results, achieving precision score of 77.1% and 79.5%, respectively, along with corresponding mAP@0.5 scores of 74.6% and 76.7%. However, an innovative YOLOv8n-seg architecture was proposed in this research, exhibiting remarkable improvements with a precision result of 83.3% and a mAP@0.5 score of 84.8%, surpassing the performance of existing YOLO variants, as shown in Figure 14. These findings highlight the efficacy of the proposed YOLOv8n-seg model, underscoring its potential for advancing object detection and segmentation tasks.

**Table 4.** Comparative results of existing approaches with the proposed approach

Reference	Model	Precision	Recall	mAP@0.5	mAP@0.5-0.95
Anbalagan, et al. [45]	YOLOv5n-seg	69.69	60.6	60.6	28.7
	YOLOv5s-seg	76.4	47.4	69.7	33.4
	YOLOv8n-seg	77.1	70.5	74.6	36.9
	YOLOv8s-seg	79.5	71.5	76.7	39.3
Islam, et al. [46]	YOLOv5	78.2	85.1	72	NA
Proposed	YOLOv8n-seg	83.3	78.2	84.8	55.7



**Figure 14.** Comparative graph of proposed modified segmentation model with existing approaches

## 5. Discussions

Detecting and classifying crops and weeds present numerous challenges that impede accurate analysis. These challenges include occlusion, where other objects obscure plants; similarity in colour and texture among different plant types; shadows cast on plants in natural light conditions; deviations in colour and texture due to illumination and lighting; and different weed species that appear similar. Additionally, inconsistencies in the appearance of the same crop or weed during different growth phases, motion blur, noise effects, and variations caused by geographical location, crop diversity, weather, and soil conditions further complicate accurate detection.

Our proposed YOLOv8n-Seg model effectively addressed several of these issues. For instance, mosaic data augmentation improved the model's robustness to occlusions and overlapping leaves by exposing it to diverse item placements and interactions. This technique enhanced the model's ability to manage occlusions better than previous methods. Furthermore, the model showed resilience to variations in lighting and shadows. However, extreme illumination conditions still pose challenges, which we plan to address through additional image preprocessing improvements.

Enhanced accuracy in detecting and classifying crops and weeds leads to more efficient weed management, reducing reliance on manual labour and herbicides. This reduces costs and minimises environmental impact, contributing to more sustainable agricultural practices. The model's robustness across various conditions enables integration into real-time monitoring systems on drones and field robots, optimising resource use and supporting better crop health assessments. These advancements contribute to more precise and eco-friendly agricultural methods.

While the model performed well across different weed species and growth phases, it faced difficulties due to regional and environmental variations. To mitigate these issues, future research will focus on the

expanding the dataset to include a broader range of settings and optimising the model to handle these diverse conditions effectively. Additionally, resource constraints impacted the model's efficiency, particularly with large datasets. To enhance real-time performance in agricultural scenarios, we are refining the YOLOv8 model for deployment on edge devices, such as autonomous drones and field robots. The model's deployment in real-world agricultural settings faces challenges due to resource requirements and calibration, necessitating computational power and field-specific adjustments. Future research should enhance training, optimise for diverse conditions, and improve computational efficiency. These revisions are intended to address the unique challenges identified and improve the model's applicability and effectiveness in practical agricultural settings.

## 6. Conclusions

This study highlights the importance of accurate segmentation in agricultural image analysis. By precisely separating crops from weeds, we can improve weed identification and support precision farming practices. Leveraging the powerful YOLOv8 object detection model, we introduced a novel methodology that integrates specialised segmentation techniques to precisely delineate individual crops and weeds, even in challenging agricultural environments. Through extensive experimentation on benchmark datasets containing 2630 diverse crop and weed species, this study highlighted the importance of accurate segmentation in agricultural image analysis. By precisely separating crops from weeds, we can improve weed identification and support precision farming practices. Our method shows encouraging performance on both 'crop' and 'weed' classes, achieving high precision, recall, mAP, and F1 score values. The robust performance of the YOLOv8 segmentation algorithm, as highlighted by its overall mAP@0.5 of 0.843 and 0.848, reaffirms its efficacy in facilitating accurate and efficient box detection and mask segmentation, respectively, in agricultural contexts. These findings emphasise the potential of our methodology to contribute significantly to advancing precision agriculture, allowing farmers to make decisions for optimised weed control and crop management strategies.

Our approach includes model pruning to reduce complexity, quantisation to lower precision, and knowledge distillation to create a smaller, more efficient model. We will also focus on hardware-specific tuning to utilise device capabilities and address real-time performance constraints. Challenges include balancing accuracy with computational efficiency and ensuring integration with edge devices.

We intend to improve the training process by including techniques like transfer learning from pre-trained models, which could hasten convergence and improve performance on our specific job. Hyperparameter optimisation, such as changes to learning rates, dropout rates, and batch sizes, will be thoroughly investigated to optimise the training process further.

Further research may explore integrating additional data sources and refining segmentation techniques to improve the algorithm's functionality in real-world agricultural practices. Additionally, research

work focuses on real-world deployment and evaluation of the developed methodology in diverse agricultural settings to assess its practical applicability and identify potential limitations.

**Declaration of Competing Interest** The authors declare that they have no known competing of interest.

## References

- [1] S. K. Gupta, S. K. Yadav, S. K. Soni, U. Shanker, and P. K. Singh, "Multiclass weed identification using semantic segmentation: An automated approach for precision agriculture," *Ecological Informatics*, vol. 78, p. 102366, 2023.
- [2] V. Balaska, Z. Adamidou, Z. Vryzas, and A. Gasteratos, "Sustainable crop protection via robotics and artificial intelligence solutions," *Machines*, vol. 11, no. 8, p. 774, 2023.
- [3] Y. Mulyani, S. Dzihan, M. A. Muhammad, and G. F. Nama, "Comparison Study of Convolutional Neural Network Architecture in Aglaonema Classification," *International Journal of Electronics and Communications System*, vol. 2, no. 2, 2022.
- [4] A. Azizi, Y. Abbaspour-Gilandeh, E. Vannier, R. Dusséaux, T. Mseri-Gundoshmian, and H. A. Moghaddam, "Semantic segmentation: A modern approach for identifying soil clods in precision farming," *Biosystems Engineering*, vol. 196, pp. 172-182, 2020.
- [5] A. Vijayakumar and S. Vairavasundaram, "Yolo-based object detection models: A review and its applications," *Multimedia Tools and Applications*, pp. 1-40, 2024.
- [6] A. Paul, R. Machavaram, D. Kumar, and H. Nagar, "Smart solutions for capsicum Harvesting: Unleashing the power of YOLO for Detection, Segmentation, growth stage Classification, Counting, and real-time mobile identification," *Computers and Electronics in Agriculture*, vol. 219, p. 108832, 2024.
- [7] Z. Luo, W. Yang, Y. Yuan, R. Gou, and X. Li, "Semantic segmentation of agricultural images: A survey," *Information Processing in Agriculture*, 2023.
- [8] A. Venkataraju, D. Arumugam, C. Stepan, R. Kiran, and T. Peters, "A review of machine learning techniques for identifying weeds in corn," *Smart Agricultural Technology*, vol. 3, p. 100102, 2023.
- [9] R. Sapkota, J. Stenger, M. Ostlie, and P. Flores, "Towards reducing chemical usage for weed control in agriculture using UAS imagery analysis and computer vision techniques," *Scientific Reports*, vol. 13, no. 1, p. 6548, 2023.
- [10] Y. Lu and S. Young, "A survey of public datasets for computer vision tasks in precision agriculture," *Computers and Electronics in Agriculture*, vol. 178, p. 105760, 2020.
- [11] R. Gerhards, D. Andujar Sanchez, P. Hamouz, G. G. Peteinatos, S. Christensen, and C. Fernandez-Quintanilla, "Advances in site-specific weed management in agriculture—A review," *Weed Research*, vol. 62, no. 2, pp. 123-133, 2022.
- [12] A. Bakhshipour, A. Jafari, S. M. Nassiri, and D. Zare, "Weed segmentation using texture features extracted from wavelet sub-images," *Biosystems Engineering*, vol. 157, pp. 1-12, 2017.
- [13] A. Perez, F. Lopez, J. Benlloch, and S. Christensen, "Colour and shape analysis techniques for weed detection in cereal fields," *Computers and electronics in agriculture*, vol. 25, no. 3, pp. 197-212, 2000.
- [14] A. Jafari, S. S. Mohtasebi, H. E. Jahromi, and M. Omid, "Weed detection in sugar beet fields using machine vision," *Int. J. Agric. Biol.*, vol. 8, no. 5, pp. 602-605, 2006.
- [15] Y. Zheng, Q. Zhu, M. Huang, Y. Guo, and J. Qin, "Maize and weed classification using color indices with support vector data description in outdoor fields," *Computers and electronics in agriculture*, vol. 141, pp. 215-222, 2017.
- [16] F. Lin, D. Zhang, Y. Huang, X. Wang, and X. Chen, "Detection of corn and weed species by the combination of spectral, shape and textural features," *Sustainability*, vol. 9, no. 8, p. 1335, 2017.
- [17] E. Hamuda, B. Mc Ginley, M. Glavin, and E. Jones, "Automatic crop detection under field conditions using the HSV

- colour space and morphological operations," *Computers and electronics in agriculture*, vol. 133, pp. 97-107, 2017.
- [18] M. F. Guerri, C. Distanto, P. Spagnolo, F. Bougourzi, and A. Taleb-Ahmed, "Deep learning techniques for hyperspectral image analysis in agriculture: A review," *ISPRS Open Journal of Photogrammetry and Remote Sensing*, p. 100062, 2024.
- [19] S. Mogilicharla and U. K. Mummadi, "The literature survey: Precision agriculture for crop yield optimization," in *AIP Conference Proceedings*, 2024, vol. 3007, no. 1: AIP Publishing.
- [20] N. Rai and X. Sun, "WeedVision: A single-stage deep learning architecture to perform weed detection and segmentation using drone-acquired images," *Computers and Electronics in Agriculture*, vol. 219, p. 108792, 2024.
- [21] R. K. Kasera, S. Gour, and T. Acharjee, "A comprehensive survey on IoT and AI based applications in different pre-harvest, during-harvest and post-harvest activities of smart agriculture," *Computers and Electronics in Agriculture*, vol. 216, p. 108522, 2024.
- [22] T. Sun *et al.*, "Weed Recognition at Soybean Seedling Stage Based on YOLOV8nGP+ NExG Algorithm," *Agronomy*, vol. 14, no. 4, p. 657, 2024.
- [23] J. Mendoza-Bernal, A. González-Vidal, and A. F. Skarmeta, "A Convolutional Neural Network approach for image-based anomaly detection in smart agriculture," *Expert Systems with Applications*, vol. 247, p. 123210, 2024.
- [24] J. You, W. Liu, and J. Lee, "A DNN-based semantic segmentation for detecting weed and crop," *Computers and Electronics in Agriculture*, vol. 178, p. 105750, 2020.
- [25] S. G. Sodjinou, V. Mohammadi, A. T. S. Mahama, and P. Gouton, "A deep semantic segmentation-based algorithm to segment crops and weeds in agronomic color images," *information processing in agriculture*, vol. 9, no. 3, pp. 355-364, 2022.
- [26] K. Osorio, A. Puerto, C. Pedraza, D. Jamaica, and L. Rodríguez, "A deep learning approach for weed detection in lettuce crops using multispectral images," *AgriEngineering*, vol. 2, no. 3, pp. 471-488, 2020.
- [27] Y. H. Kim and K. R. Park, "MTS-CNN: Multi-task semantic segmentation-convolutional neural network for detecting crops and weeds," *Computers and Electronics in Agriculture*, vol. 199, p. 107146, 2022.
- [28] M. Fawakherji, A. Youssef, D. Bloisi, A. Pretto, and D. Nardi, "Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation," in *2019 third IEEE international conference on robotic computing (IRC)*, 2019, pp. 146-152: IEEE.
- [29] 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.
- [30] A. Khan, T. Ilyas, M. Umraiz, Z. I. Mannan, and H. Kim, "Ced-net: crops and weeds segmentation for smart farming using a small cascaded encoder-decoder architecture," *Electronics*, vol. 9, no. 10, p. 1602, 2020.
- [31] B. B. Sapkota *et al.*, "Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton," *Scientific Reports*, vol. 12, no. 1, p. 19580, 2022.
- [32] R. Sapkota, D. Ahmed, and M. Karkee, "Comparing YOLOv8 and Mask RCNN for object segmentation in complex orchard environments," *arXiv preprint arXiv:2312.07935*, 2023.
- [33] J. Cui, F. Tan, N. Bai, and Y. Fu, "Improving U-net network for semantic segmentation of corns and weeds during corn seedling stage in field," *Frontiers in Plant Science*, vol. 15, p. 1344958, 2024.
- [34] 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.
- [35] L. Hashemi-Beni, A. Gebrehiwot, A. Karimodini, A. Shahbazi, and F. Dorbu, "Deep convolutional neural networks for weeds and crops discrimination from UAS imagery," *Frontiers in Remote Sensing*, vol. 3, p. 755939, 2022.
- [36] H. Zhang *et al.*, "Automated delineation of agricultural field boundaries from Sentinel-2 images using recurrent residual U-Net," *International Journal of Applied Earth Observation and Geoinformation*, vol. 105, p. 102557, 2021.
- [37] D. Keerthana, V. Venugopal, M. K. Nath, and M. Mishra, "Hybrid convolutional neural networks with SVM classifier



- for classification of skin cancer," *Biomedical Engineering Advances*, vol. 5, p. 100069, 2023.
- [38] D. Vijayalakshmi, P. Elangovan, and M. K. Nath, "METHODOLOGY FOR IMPROVING DEEP LEARNING-BASED CLASSIFICATION FOR CT SCAN COVID-19 IMAGES," *Biomedical Engineering: Applications, Basis and Communications*, p. 2450008, 2024.
- [39] P. Elangovan, D. Vijayalakshmi, and M. K. Nath, "Covid-19net: An effective and robust approach for covid-19 detection using ensemble of convnet-24 and customized pre-trained models," *Circuits, Systems, and Signal Processing*, vol. 43, no. 4, pp. 2385-2408, 2024.
- [40] Roboflow Platform -Weed detection Computer Vision Project. (2023). Available: <https://universe.roboflow.com/koley/weed-detection-535r5>
- [41] C. Dewi, R.-C. Chen, Y.-C. Zhuang, X. Jiang, and H. Yu, "Recognizing road surface traffic signs based on YOLO models considering image flips," *Big data and cognitive computing*, vol. 7, no. 1, p. 54, 2023.
- [42] A. Mehmood, M. Ahmad, and Q. M. Ilyas, "On precision agriculture: enhanced automated fruit disease identification and classification using a new ensemble classification method," *Agriculture*, vol. 13, no. 2, p. 500, 2023.
- [43] X. Yue, K. Qi, F. Yang, X. Na, Y. Liu, and C. Liu, "RSR-YOLO: a real-time method for small target tomato detection based on improved YOLOv8 network," *Discover Applied Sciences*, vol. 6, no. 5, p. 268, 2024.
- [44] G. Yang, J. Wang, Z. Nie, H. Yang, and S. Yu, "A lightweight YOLOv8 tomato detection algorithm combining feature enhancement and attention," *Agronomy*, vol. 13, no. 7, p. 1824, 2023.
- [45] T. Anbalagan, M. K. Nath, D. Vijayalakshmi, and A. Anbalagan, "Analysis of various techniques for ECG signal in healthcare, past, present, and future," *Biomedical Engineering Advances*, vol. 6, p. 100089, 2023.
- [46] A. Islam, S. R. S. Raisa, N. H. Khan, and A. I. Rifat, "A deep learning approach for classification and segmentation of leafy vegetables and diseases," in *2023 International Conference on Next-Generation Computing, IoT and Machine Learning (NCIM)*, 2023, pp. 1-6: IEEE.