ORIGINAL RESEARCH Open access

Deep Learning-Based Weed and Crop Detection for Precision Agriculture Using Ground Robotic Systems

Daniel Kim[1*], Jihoon Park[1], Min-seo Kang[2], Hyun-woo Jung[3]

Abstract

Deploying a machine-vision framework for on-the-fly, site-specific weed control marks a pivotal advance toward the creation of intelligent spraying robots. An intelligent vision system, engineered through deep learning-based object detectors, can reliably distinguish weeds from crops in imagery. The data-centric and model-centric pathways for crafting deep learning solutions each confer distinct benefits, contingent upon environmental and non-environmental variables. To gauge the efficacy of real-field weed identification across eight crop types, Yolov8, Yolov9, and a bespoke Yolov9 architecture were trained and benchmarked on RGB images sourced from four sites (Casselton, Fargo, and Carrington) over two years in the U.S. Great Plains. The study covered eight crops—dry bean, canola, corn, field pea, flax, lentil, soybean, and sugar beet—alongside five weed species—horseweed, kochia, redroot pigweed, common ragweed, and water hemp. Six Yolov8 and eight Yolov9 model configurations were trained on annotated weed/crop snapshots from the four individual sites and a pooled multisite compilation. Detection capability was quantified via mean average precision (mAP50) scores across five datasets, eight crops, and five weeds. Outcomes from the detection assessment yielded lofty mAP50 figures, hitting 86.2%. Species-level mAP50 spanned from 80.8% to 98%. Findings underscored that model behavior shifted according to the architecture variant (model-centric), the geographical setting and associated environment (data-centric), the volume of training instances (data-centric), annotation fidelity (data-centric), and target dimensions within the frame (data-centric). Even so, the streamlined, customized Yolov9 variant is poised to play a vital role in assembling a live, vision-driven precision weed management pipeline.

Keywords Deep learning, Object detection, YOLO, Data-centric, Model-centric, Weed detection

*Correspondence:
Daniel Kim
daniel.kim@khu.ac.kr
[1] Department of AI for Pear and Apple Blossom Thinning Prediction, College of Agriculture, Kyung Hee University, Seoul, South Korea
[2] Department of Smart Fruit Bagging and Color Analysis, College of Agriculture, Chung-Ang University, Seoul, South Korea
[3] Department of Machine Learning for Frost Damage Forecasting, College of Agriculture, Kangwon National University, Chuncheon, South Korea

Introduction

Smart farming refers to the integration of cutting-edge technologies—sensors, big data streams, Internet of Things (IoT) frameworks, and artificial intelligence—to enhance agricultural output and ecological stewardship. These technology-driven methods can simultaneously increase food production volume and quality while reducing harmful environmental impacts. Moreover, smart farming aims to engineer precise, self-governing systems that perform field operations [1]. The rise of deep learning in artificial intelligence (AI) has paved the way for intelligent machine vision modules for tasks such as fruit sizing and picking [2, 3], pest suppression [4], pathology diagnosis [5], and weed abatement [6]. Vision-based systems hold promise across a spectrum of agricultural functions, including robotic traversal of fields and pinpoint

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

identification and geolocation of entities (crops, fruits, pests, and weeds) [1, 7]. Such fine-grained detection of weeds and crops propels progress in spatially selective weed management tactics.

Spatially selective weed control forms an indispensable pillar of smart farming, targeting weed suppression within crop stands. Robust weed governance curtails yield losses, enhances both produce quality and volume, and can reduce farm input expenditures [8]. Discerning weeds and crops in digital images constitutes the foundational, inaugural stage toward building a real-time, site-responsive weed-mitigation apparatus within precision weed management. Models capable of concurrently flagging both crops and weeds grant considerable operational latitude. For example, via downstream image processing informed by crop detections, these models can flag weed species absent from the training corpus. This entails masking recognized crop zones and subsequently scanning residual pixel regions for weedy growth.

For live deployment, weed-detection algorithms must be computationally efficient when running on embedded hardware aboard weed-control robots. In the deep learning realm, single-pass object detectors like You Only Look Once (YOLO) and Single Shot Detector (SSD) [9, 10] are notably cheaper to compute and thus preferred over two-stage counterparts [11]. The Yolo family [12-14], in particular, has achieved widespread adoption for weed spotting and localization, thanks to its high accuracy and fast inference [15-18].

Notwithstanding Yolo's merits in weed identification, recent investigations have largely remained restricted to monoculture settings for weed species recognition [10, 15, 17, 19]. Furthermore, Yolo models demand training on explicit weed and crop classes for field deployment. Yet, the literature omits several pervasive weed and crop taxa typical of the northern Great Plains region [9, 16, 20-23]. Thus, an inquiry into the relative performance of site-tuned models versus an aggregated, multi-location model could prove fruitful for advancing precision weed control tailored to this geographic area.

In addition, most prior efforts focus on single-location data corpora, with only a handful of works addressing multi-location data. For instance, the studies by [5, 16, 18, 22, 24, 25] relied exclusively on imagery from a single site. By contrast, investigations by Refs. [9, 10, 26] leveraged multi-location data to develop a unified weed-detection model, omitting site-specific analogues. Precision weed management could derive tangible gains from developing location-tuned models. Hypothetically, if a deep learning model trained exclusively on location B imagery (Model B) outperforms models trained on location A or a fusion set in detecting waterhemp, it could be directly embedded into a robotic rig for focused waterhemp eradication. This data-oriented philosophy in model construction and fielding could directly address a grower's most troublesome weed flora.

Beyond locality, the bulk of earlier work tackles weed/crop detection as binary classification with localization, with sparse attention devoted to multi-class differentiation encompassing weed and crop species [9, 10, 16, 17, 24-26], albeit with notable outliers [27-30]. The present investigation advances multi-species weed and crop detection by leveraging deep neural architectures to discriminate a broader spectrum of cultivated and weedy plants. Forging models capable of recognizing an expanded catalog of weed and crop types is critical, as it illuminates how algorithmic performance varies across diverse plant morphologies and enables their use in heterogeneous landscapes where multiple crops coexist.

To address the gaps enumerated above, Yolov8, Yolov9, and a tailored, lightweight Yolov9 weed-detection model were trained on imagery collected at four field sites in North Dakota, United States. First, the Yolov8 and Yolov9 model iterations formulated in this work can classify eight crop species (Dry Bean, Canola, Corn, Field Pea, Flax, Lentil, Soybean, and Sugar Beet) and five weed species (Horseweed, Kochia, Redroot Pigweed, Common Ragweed, and Waterhemp). Second, models were derived from each location's dataset in isolation, alongside a merged multisource compilation (comprehensive model), to pinpoint the optimal configuration. Third, the strongest- and weakest-performing weed and crop classes were catalogued for the developed Yolov8 and Yolov9 versions. This research may furnish the groundwork for crafting intelligent sprayers dedicated to site-specific precision weed management across eight economically important U.S. cropping systems.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Materials and Methods

Experiment field location and design

Field trials were executed over two growing seasons at four North Dakota locales, USA: Fargo (Grand Farm Test Site), Carrington, and a pair of sites in Casselton. Cultivated crop species comprised Dry Bean, Canola, Corn, Field Pea, Flax, Lentil, Soybean, and Sugar Beet **(Figure 1)**; weed species included Horseweed, Kochia, Redroot Pigweed, Common Ragweed, and Waterhemp. The 2021 plot configuration in Casselton and Carrington followed a randomized layout, as depicted in **Figure 2**, with 32 total plots—four replicates per crop species. The 2022 layout, situated in Fargo and Casselton, diverged from the prior year's scheme, featuring three complete replicates, each containing eight plots assigned to the eight crop species **(Figure 3)**. The five weed species were propagated within greenhouses at Fargo and Carrington before being transplanted into inter-row zones during both the 2021 and 2022 seasons. A controlled greenhouse nursery was used for weed rearing because it was impractical to secure a fully representative weed cohort under natural field conditions. Inter-row spacing was wider in the 2021 plots (Lentil, Field Pea, and Flax in **Figure 1**) than in the 2022 plots (Dry Bean, Canola, Corn, Soybean, and Sugar Beet in **Figure 1**), due to data-acquisition hardware constraints.

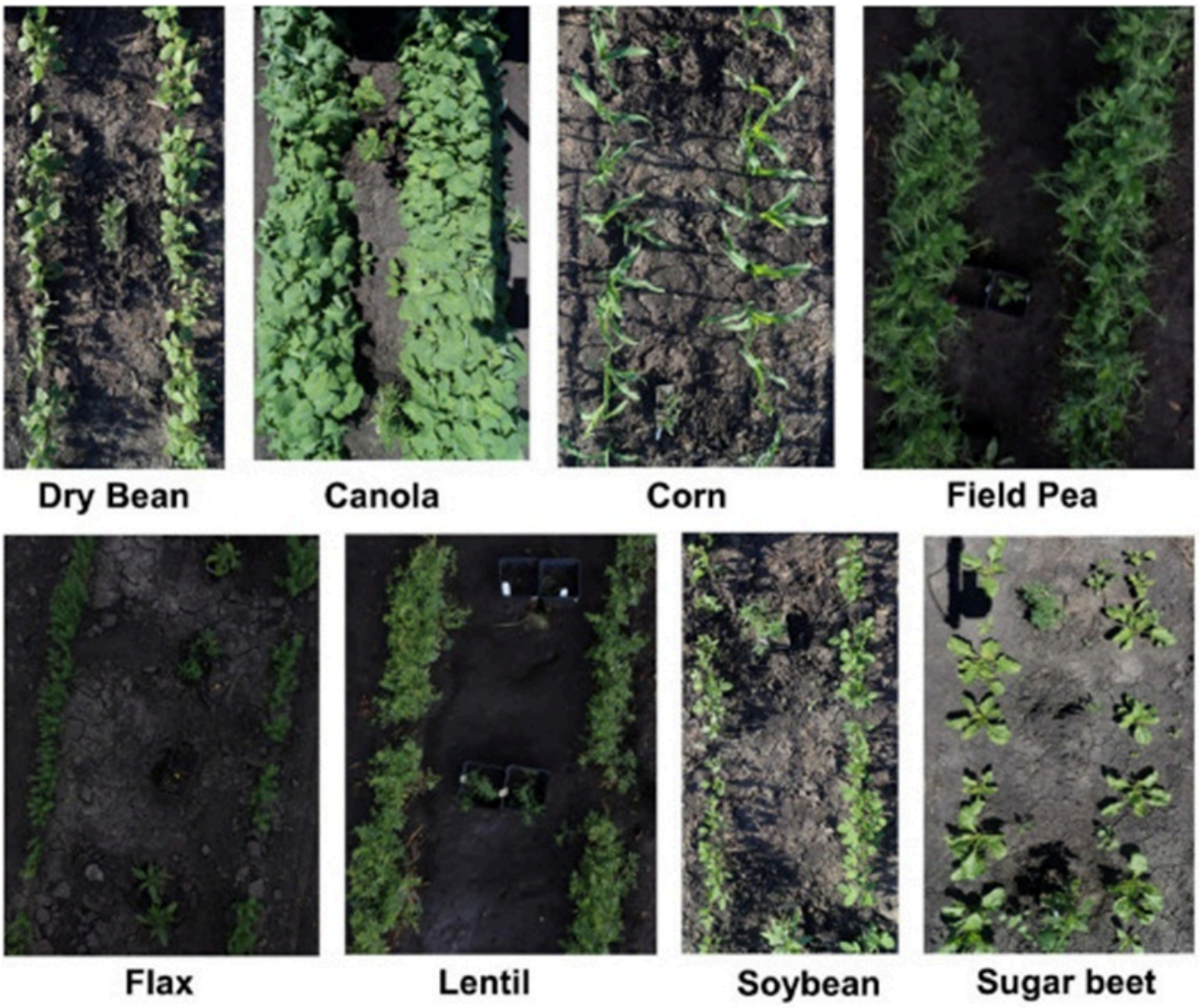


Figure 1. Eight crop species were planted in the field, with five weed species transplanted in between the rows. The lentil and Field Pea rows image is from Carrington 2021, the Flax rows image is from Casselton 2021, the Canola and Sugar

© 2024 The Author(s). | Published by Agro Intelligence Research Press

beet rows image is from Casselton 2022, and the Dry Bean, Corn, and Soybean rows image is from Fargo 2022.

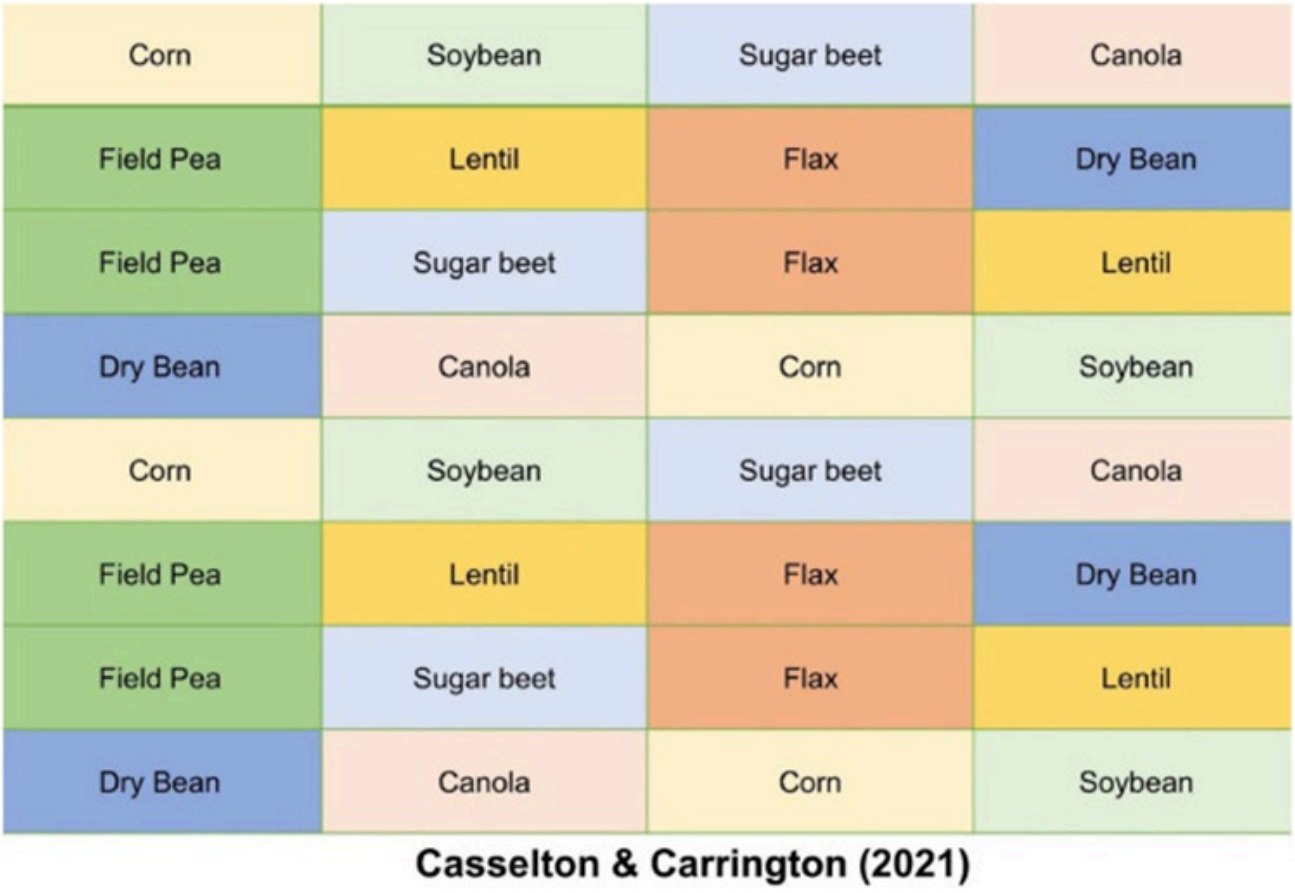


Figure 2. Field experiment design for 2021, where each box represents crop rows with weeds in between the rows.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

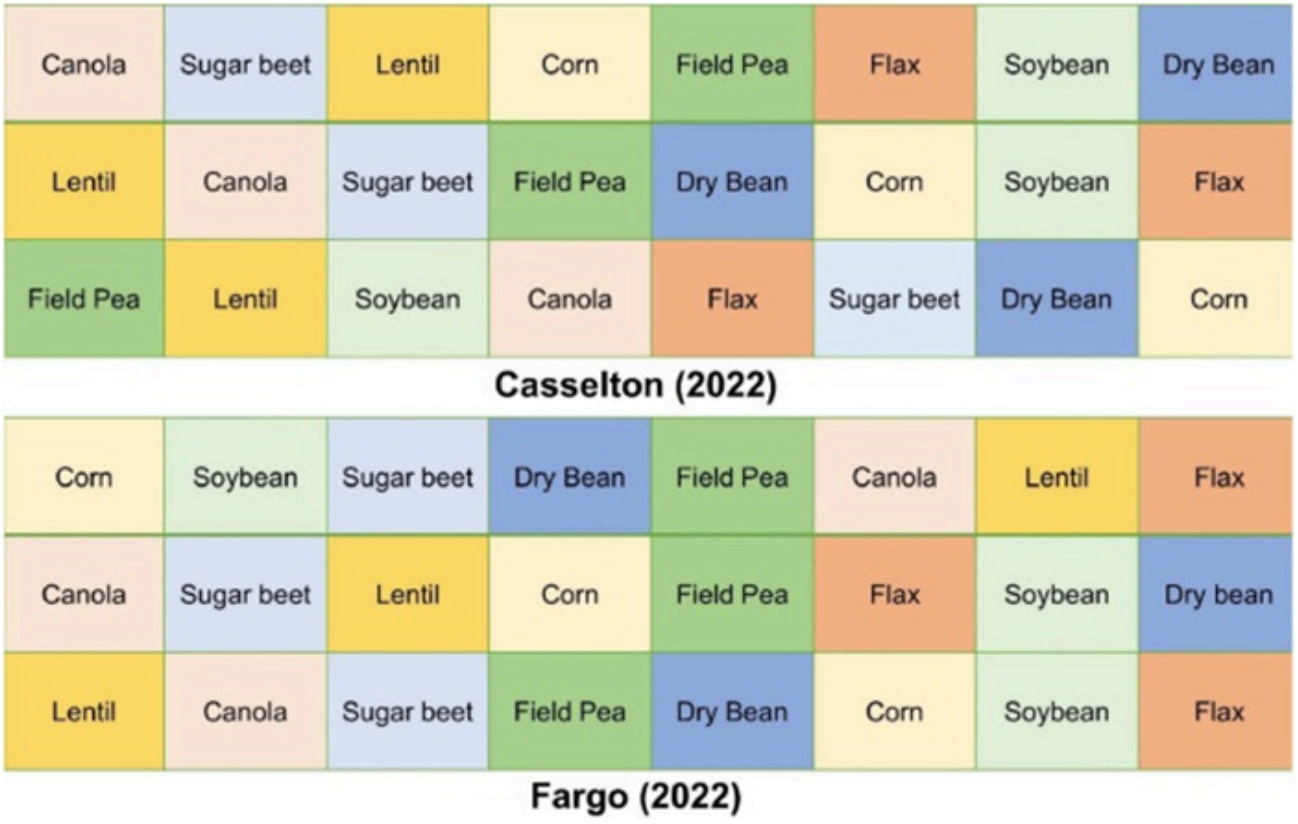


Figure 3. Field experiment design for 2022, where each box represents crop rows with weeds in between the rows.

Data acquisition

Image collection was performed via two distinct ground-based, high-throughput phenotyping conveyances—dubbed Ground Robots—during the 2021 and 2022 seasons. In 2021, the Weedbot gathered weed and crop imagery. In contrast, in 2022, the Mini Weedbot fulfilled that role, with both platforms shown in **Figure 4**. A Canon EOS 90D digital camera (Canon Inc., Tokyo, Japan) was mounted on each system **(Figure 4)**, which could be adjusted to varying heights. The camera was positioned at 48.5±2.5 inches for the Weedbot and 48.5 inches for the MiniWeedbot, guaranteeing a field of view that encompassed at least two crop rows **(Figure 4)**. The camera was linked to a controller configured for a 1-second interval, and acquisition commenced once the Weedbot and MiniWeedbot were navigated via joystick. **Figure 1** displays the resultant weed and crop imagery captured by the Weedbot (Lentil, Field Pea, and Flax) and the Mini Weedbot (Dry Bean, Corn, Canola, Soybean, and Sugar Beet).

© 2024 The Author(s). | Published by Agro Intelligence Research Press

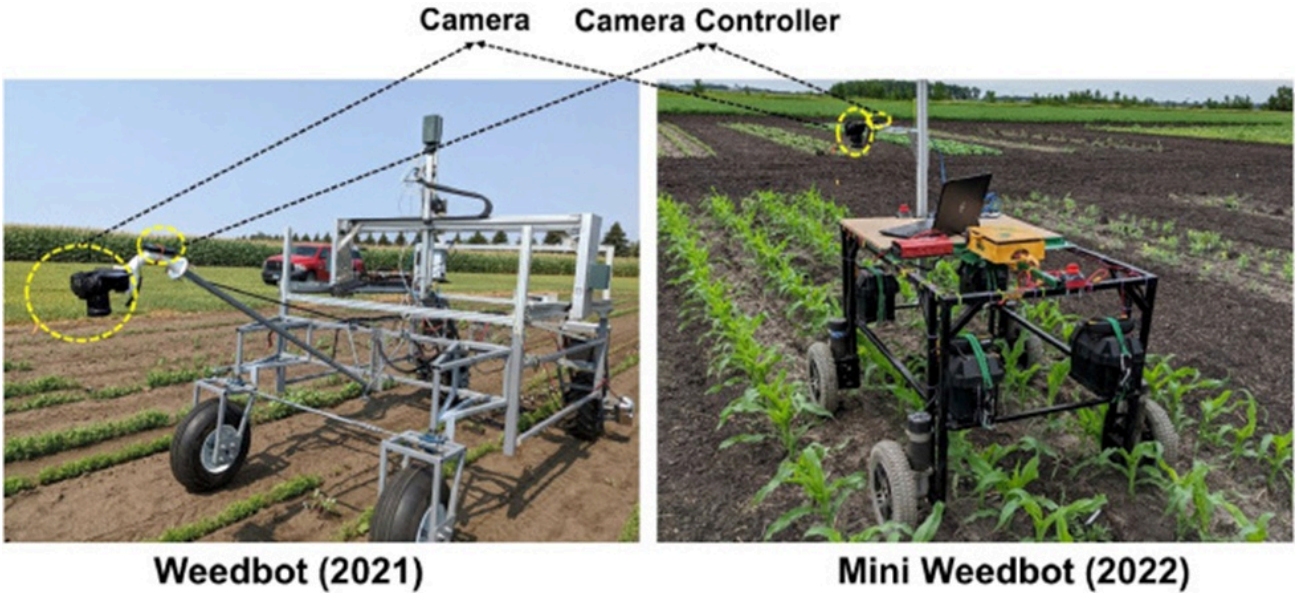


Figure 4. Two high-throughput phenotyping systems (Ground Robots), the Weedbot and the Mini Weedbot, acquiring field images from Casselton and Carrington in 2021 and 2022, respectively.

Data cleaning and data labeling

Cleaning the raw data constitutes an essential foundation for building a dependable deep learning model. The collected weed and crop footage contained irrelevant and repetitive frames—situations in which the camera veered off the planted rows or the ground platform remained stationary. These unhelpful and duplicate captures were stripped away to improve overall dataset hygiene. Furthermore, the Canon EOS camera's onboard storage occasionally produced identical file names for images residing in different directories; a suite of Python scripts was employed to systematically rename the files and avert collisions. Once renamed, every image underwent manual cropping to discard extraneous pixels outside the cultivated rows, thereby reducing file size and shortening subsequent training cycles.

Following data cleaning, the next stage involved annotating weed and crop objects under field conditions using the LabelImg toolkit [31]. **Table 1** tallies the bounding-box instance counts for each weed and crop class across the Casselton, Carrington, and Fargo collections. The overall image count reached 2024, 1700, 358, and 751 for Casselton 2021, Carrington 2021, Casselton 2022, and Fargo 2022, respectively. Casselton 2021 generated the richest image pool, while Casselton 2022 yielded the sparsest. Surprisingly, the fewest annotated objects appeared in the Fargo 2022 corpus. This partly reflects weak crop establishment in some plots, alongside oversized bounding boxes necessitated by heavy inter-row canopy closure. Inspecting object frequencies across all four sources reveals stark imbalances, rooted in the fact that germination timing and seedling vigor for the five weed and eight crop species fluctuate considerably with shifting environmental drivers—temperature, humidity, and soil moisture [32-36]. These fluctuating conditions triggered uneven emergence and growth trajectories, with some seeds never sprouting.

Table 1. The number of each weed and crop species inside the bounding box after labeling the images from all four locations.

Weed/Crop	Casselton 2022	Fargo 2022	Carrington 2021	Casselton 2021
Dry Bean	999	829	1116	2537
Canola	782	506	638	1392
Corn	729	383	736	1625

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Field Pea	630	403	2188	1433
Flax	473	940	1840	2875
Horseweed	219	110	127	553
Kochia	239	106	692	1739
Lentil	519	423	2180	1453
Common Ragweed	496	106	520	2471
Redroot Pigweed	132	57	683	1208
Soybean	1546	747	1806	1873
Sugar beet	591	243	882	228
Waterhemp	282	113	334	1371
Total object	**7637**	**4966**	**13742**	**20758**
Total images	**358**	**751**	**1700**	**2024**

Yolo model training with weed and crop images

The You Only Look Once (YOLO) framework was first proposed by Redmon *et al.* [12] in their seminal computer vision paper, deliberately architected for real-time operation via a single-stage detection pipeline. That foundational publication by Redmon *et al.* [12] lays out the system's internal mechanics in detail. Since that debut, Yolo has undergone successive generational upgrades, each lifting performance through structural redesign. The two Yolo real-time detectors used in the present work were Yolov8 [37] and Yolov9 [38]. Their adoption as the chief backbones for experimentation was motivated by their leading-edge results on publicly available benchmarks. The end goal is to eventually port the trained model onto a robotic carrier currently running YOLOv4 for on-the-fly weed recognition [39]. Advancing beyond prior Yolo lineages, Yolov8 brings fresh design concepts and efficiency tweaks. Major improvements include a reworked backbone-and-neck topology, an anchor-free split Ultralytics detection head, a carefully tuned accuracy–speed trade-off, and a broad palette of pretrained weights [37]. Exhaustive architectural details for Yolov8 can be consulted in the review authored by Sohan *et al.* [40]. **Table 2** lists the specifications of six Yolov8 model variants, covering parameter volumes, GFLOPs, and on-disk size. The investigation compared both compact and high-capacity Yolov8 configurations.

Table 2. Number of parameters, model size, and GFLOPs in the Yolo-v8 and Yolo-v9 models' variants. CW refers to custom weed models (cw1, cw2, cw3, cw4, cw5). The lightest Yolov8 and Yolov9 models are indicated by both bold and underscored parameter values, while the bold values correspond to the largest models.

Model name	Model size (MB)	GFLOPs	Number of parameters
Yolov8l	85.56	164.9	43616631
Yolov8m	50.78	78.7	25847287
Yolov8n	6.06	8.1	**3008183**
Yolov8s	21.96	28.5	11130615
Yolov8x	133.49	257.5	68136087
Yolov8 × 6	190.57	260.7	**97296324**
Yolov9	119.50	207.1	47355310

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Yolov9c	100.40	236.8	50725998
Yolov9e	136.68	240.8	**68566318**
Yolov9_cw1	109.50	246.3	55398382
Yolov9_cw2	108.632	242.8	54956014
Yolov9_cw3	93.60	207.1	47355310
Yolov9_cw4	60.2	149.4	30392356
Yolov9_cw5	10.5	26.5	**5044388**

In contrast to the Yolov8 arm, where only unmodified factory variants were trialed, this work deployed both stock Yolov9 models (Yolov9, Yolov9e, Yolov9c) and purpose-built custom derivatives. In concrete terms, three off-the-shelf Yolov9 models and five tailored variants were benchmarked **(Table 2)**. Yolov9 signaled a significant advance for real-time object detection by introducing Programmable Gradient Information (PGI) alongside the Generalized Efficient Layer Aggregation Network (GELAN) [38]. Within this study, the standard Yolov9 skeleton was refactored to produce Yolov9_CW1, Yolov9_CW2, Yolov9_CW3, Yolov9_CW4, and Yolov9_CW5. The customization drive targeted efficiency by dialing down convolution filter parameters within the feature extraction chain. This tactic shrank the total parameter budget, yielding a lightweight model series. In **Table 2**, the steadily declining parameter tallies from Yolov9_cw1 down to Yolov9_cw5 mirror the stepwise trimming of convolution filter values. Modifications in filter quantities between the stock Yolov9 and the customized versions within the backbone and neck segments are enumerated in **Tables 3 and 4**, respectively. The default configuration settings for Yolov9, retrieved from the official GitHub repository (https://github.com/WongKinYiu/Yolov9), served as the starting blueprint for engineering the lightweight customized variants.

Table 3. Comparison of backbone layers and parameters in Yolov9 and customized lightweight Yolov9 models.

Layer name	**Yolov9 and customize Yolov9 models' parameters**					
Empty Cell	**Yolov9-CW5**	**Yolov9-Cw4**	**Yolov9-CW3**	**Yolov9-CW2**	**Yolov9-CW1**	**Yolov9**
Silence[a]	[]	[]	[]	[]	[]	[]
Conv[b]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[64, 3, 2]	[64, 3, 2]
Conv	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[64, 3, 2]	[128, 3, 2]
RepNCSPELAN4[c]	[64,32,16,1]	[64,32,16,1]	[64,32,16,1]	[256,128,64,1]	[256,128,64,1]	[256,128,64,1]
Conv	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[64, 3, 2]	[256, 3, 2]
RepNCSPELAN4	[64,32,16,1]	[64,32,16,1]	[64,32,16,1]	[512,256,28,1]	[512,256,128,1]	[512,256,128,1]
Conv	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[64, 3, 2]	[512, 3, 2]
RepNCSPELAN4	[64,64,32,1]	[64,64,32,1]	[64,64,32,1]	[512,512,256,1]	[512,512,256,1]	[512,512,256,1]
Conv	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[32, 3, 2]	[64, 3, 2]	[512, 3, 2]
RepNCSPELAN4	[128, 128, 64, 1]	[128, 128, 64, 1]	[128,128,64,1]	[512,512,256,1]	[512,512,256,1]	[512,512,256,1]

[a]: This layer does not perform any operation on data.

[b]: Convolution layer.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

[c]: Layer that combines elements from CSPNet (Cross Stage Partial Networks) and ELAN (Efficient Layer Aggregation Network).

Table 4. Comparison of neck region layers and parameters in Yolov9 and customized lightweight Yolov9 models.

Layer name	**Yolov9 and customize Yolov9 models' parameters**					
Empty Cell	**Yolov9-CW5**	**Yolov9-Cw4**	**Yolov9-CW3**	**Yolov9-CW2**	**Yolov9-CW1**	**Yolov9**
SPPELAN[a]	[256, 128]	[512, 256]	[512, 256]	[512, 256]	[512, 256]	[512, 256]
Upsample[b]	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']
Concat[c]	[[−1, 7]]	[[−1, 7]]	[[−1, 7]]	[[−1, 7]]	[[−1, 7]]	[[−1, 7]]
RepNCSPELAN4	[128, 128, 64, 1]	[128, 128, 64, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]
Upsample	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']	[None, 2, 'nearest']
Concat	[[−1, 5]]	[[−1, 5]]	[[−1, 5]]	[[−1, 5]]	[[−1, 5]]	[[−1, 5]]
RepNCSPELAN4	[1, 64, 32, 1]	[1, 64, 32, 1]	[256, 256, 128, 1]	[256, 256, 128, 1]	[256, 256, 128, 1]	[256, 256, 128, 1]
Conv	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]
Concat	[[−1, 13]]	[[−1, 13]]	[[−1, 13]]	[[−1, 13]]	[[−1, 13]]	[[−1, 13]]
RepNCSPELAN4	[64, 64, 32, 1]	[64, 64, 32, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]
Conv	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]
Concat	[[−1, 10]]	[[−1, 10]]	[[−1, 10]]	[[−1, 10]]	[[−1, 10]]	[[−1, 10]]
RepNCSPELAN4	[128, 128, 64, 1]	[128, 128, 64, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]
CBLinear[d]	[[256]]	[[256]]	[[256]]	[[256]]	[[256]]	[[256]]
CBLinear	[[256, 512]]	[[256, 512]]	[[256, 512]]	[[256, 512]]	[[256, 512]]	[[256, 512]]
CBLinear	[[256, 512, 512]]	[[256, 512, 512]]	[[256, 512, 512]]	[[256, 512, 512]]	[[256, 512, 512]]	[[256, 512, 512]]
Conv	[64, 3, 2]	[64, 3, 2]	[64, 3, 2]	[64, 3, 2]	[64, 3, 2]	[64, 3, 2]
Conv	[128, 3, 2]	[128, 3, 2]	[128, 3, 2]	[128, 3, 2]	[128, 3, 2]	[128, 3, 2]
RepNCSPELAN4	[128, 64, 32, 1]	[256, 128, 64, 1]	[256, 128, 64, 1]	[256, 128, 64, 1]	[256, 128, 64, 1]	[256, 128, 64, 1]
Conv	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]	[256, 3, 2]
CBFuse	[[23, 24, 25, −1,0, 0, 0]]	[[23, 24, 25, −1,0, 0, 0]]	[[23, 24, 25, −1,0, 0, 0]]	[[23, 24, 25, −1,0, 0, 0]]	[[23, 24, 25, −1,0, 0, 0]]	[[23, 24, 25, −1,0, 0, 0]]

© 2024 The Author(s). | Published by Agro Intelligence Research Press

RepNCSPELAN4	[128, 64, 32, 1]	[512, 256, 128, 1]	[512, 256, 128, 1]	[512, 256, 128, 1]	[512, 256, 128, 1]	[512, 256, 128, 1]
Conv	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]
CBFuse[e]	[[24, 25, −1,1, 1]]	[[24, 25, −1,1, 1]]	[[24, 25, −1,1, 1]]	[[24, 25, −1,1, 1]]	[[24, 25, −1,1, 1]]	[[24, 25, −1,1, 1]]
RepNCSPELAN4	[128, 128, 64, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]
Conv	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]	[512, 3, 2]
CBFuse	[25, −1,2]	[25, −1,2]	[25, −1,2]	[25, −1,2]	[25, −1,2]	[25, −1,2]
RepNCSPELAN4	[128, 128, 64, 1]	[512, 512, 256, 1]	[512, 512, 256, 1]	[512,512,256, 1]	[512,512,256, 1]	[512,512,256, 1]

[a]: Layer that combines elements of Spatial Pyramid Pooling (SPP) and Efficient Layer Aggregation Network (ELAN).

[b]: Layer that increases spatial resolution.

[c]: Layer to merge multiple tensor inputs.

[d]: Convolutional layer that splits the output into multiple channels.

[e]: Layer that fuses multiple tensors by interpolating and summing them.

In the experimental design, models were fitted to each location's dataset in isolation and to a single pooled dataset combining all four sites, enabling direct cross-comparison. To prepare the weed-and-crop dataset for Yolo training, all images were first split into training, validation, and test partitions at a 70:15:15 ratio for each location. The test partition was held back exclusively for evaluating generalization across sites. Training and evaluation cycles ran on a Lambda computing workstation (Lambda Labs, San Jose, USA) with four NVIDIA graphics processing units (GPUs), each with 24248 megabytes (MB) of memory.

Yolo model performance evaluation

The Yolo models' performance was measured using average precision (AP) and mean average precision (mAP) [11, 41-43]. Calculating AP and mAP first requires establishing true positive (TP), false positive (FP), and false negative (FN) counts. In object detection, the intersection over union (IoU) metric quantifies the spatial overlap between a predicted box and the ground truth. **Figure 5** visualizes the IoU concept, where rectangles R1 and R2 denote the ground-truth annotation and the model's prediction, respectively. The IoU threshold (α) was set to 50% to compute the mAP50 metric. A TP is registered when the model outputs a positive detection, the ground truth is positive, and the IoU meets or surpasses the threshold (IoU ≥ α). An FP is logged when the model outputs a positive detection while either the ground truth is negative or the IoU falls short of the threshold (IoU < α). An FN occurs when a ground-truth object is completely missed by the detector. Precision and recall are computed from TP, FP, and FN as given in **Eqs. 1 and 2**, respectively. These two quantities then underpin the calculation of average precision (AP) and mean average precision (mAP), as formulated in **Eqs. 3 and 4**. In the AP expression **(Eq. 3)**, Rn and Pn correspond to the recall and precision at the n-th threshold, whereas N in the mAP expression **(Eq. 4)** indicates the total class count.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

$$Precision = \frac{Predicted\ area\ in\ ground\ truth}{Total\ area\ of\ predicted\ region} = \frac{TP}{TP+FP} \quad (1)$$

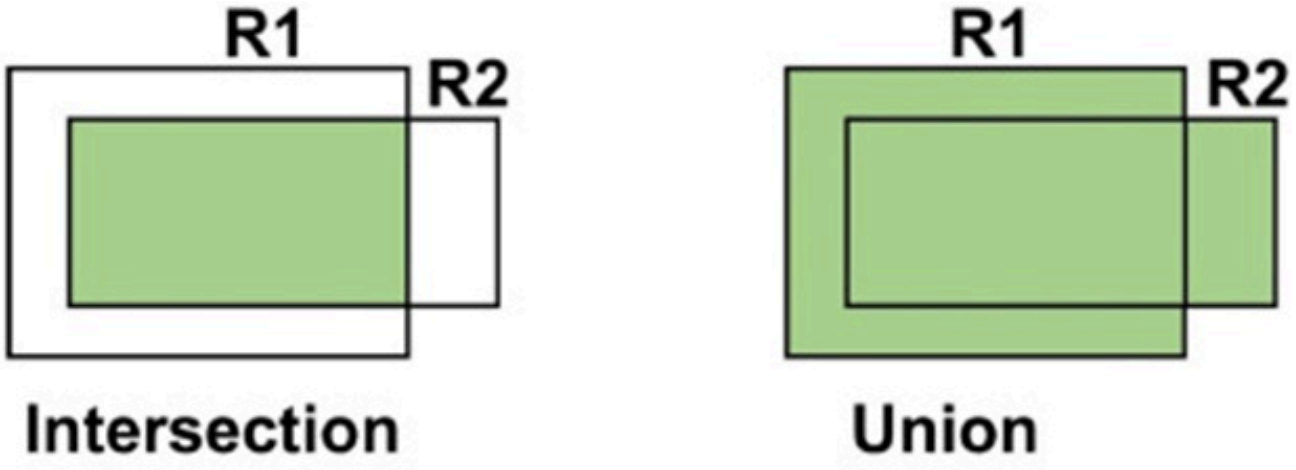
$$Recall = \frac{Ground\ truth\ area\ in\ predicted\ region}{Total\ area\ of\ ground\ truth\ region} = \frac{TP}{TP+FN}$$

(2)

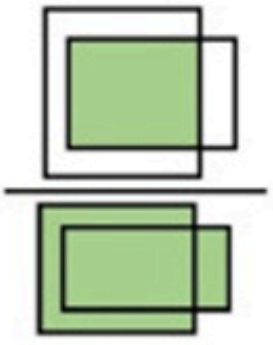
$$AP = \sum_n (R_n - R_{n-1})\, P_n \quad (3)$$

$$Mean\ average\ precision\ (mAP) = \frac{1}{N} \sum_{i=1}^{N} AP_i$$

(4)



$$IoU = \frac{R1 \cap R2}{R1 \cup R2} =$$



Intersection over Union (IoU)

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Figure 5. Illustration of intersection over union (IoU).

Results and Discussion

Yolo model comparison based on four locations

The detection capability of the Yolov8 and Yolov9 families was compared after training on five separate data compositions (Carrington 2021, Carrington 2022, Casselton 2021, Fargo 2022, and a combined pool). **Figure 6** charts the mAP50 (mean average precision at 50% Intersection over Union) for six Yolov8 and eight Yolov9 variants across these five datasets. Among the location-specific sets, models trained solely on Carrington 2022 posted the weakest results. Carrington 2021 and Casselton 2021 data delivered considerably stronger performance relative to the 2022 collections (Fargo 2022, Carrington 2022). This discrepancy likely originates from the smaller image reservoirs available for Carrington 2022 and Fargo 2022 during training.

Furthermore, constrained sample sizes tend to curtail visual diversity, eroding the data quality essential for model generalization [44]. When imagery from all four locations was merged into a combined dataset, performance rose above that seen on Fargo 2022 and Carrington 2022 alone, yet still trailed the figures achieved with Carrington 2021 and Casselton 2021. The mAP50 readings exposed a tangible gap between the combined model or models trained on 2021 data versus those trained on 2022 data. These patterns reinforce the notion that model success is as much a function of the training data as of the network design itself. Detectors exposed to broader visual variation typically generalize more reliably, echoing observations from a survey paper [14]. In this study, the 2021 imagery consisted of captures taken from a single pass direction along the rows, whereas the 2022 imagery included both forward and return passes. These findings align with the growing momentum in the research community behind data-centric AI development.

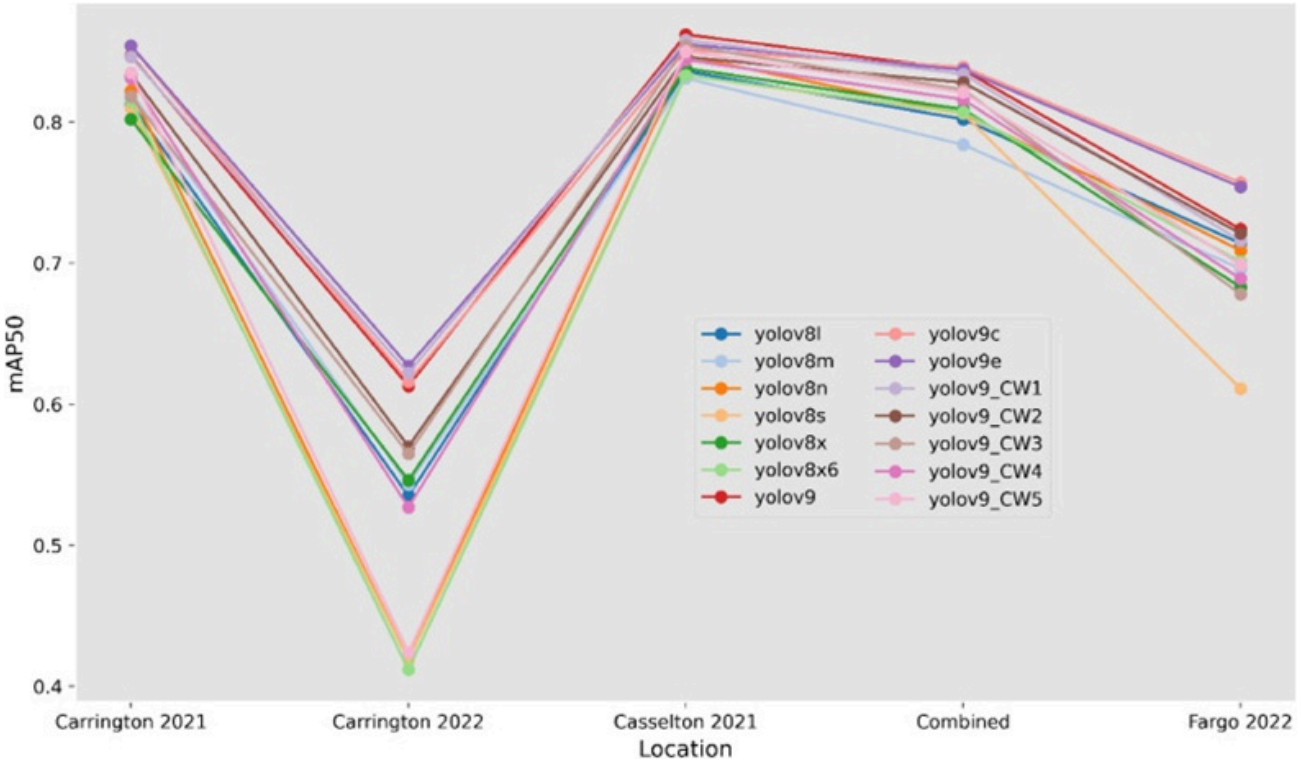


Figure 6. The performance of six Yolov8 and eight Yolov9 models on five different datasets. Five Yolov9 variants with the suffix _CWi (i=1,2,3,4,5) were customized models.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Figure 6 also uncovers performance fluctuations among the six Yolov8 and eight Yolov9 model variants. Visual inspection finds no consistent trend. To illustrate, Yolov9e came out on top for Carrington data in both 2021 and 2022. However, the lead position shifted for Casselton 2021 (Yolov9), the combined set (Yolov9c), and Fargo 2022 (Yolov9c). **Table 5** records the mAP50 scores for Yolov8 and Yolov9 models on the four separate location sets and the combined compilation. Broadly speaking, Yolov9 models outperformed their Yolov8 equivalents across all five datasets, though exceptions occurred; in selected cases, particular Yolov8 variants edged out their Yolov9 rivals. These performance swings can be traced to architectural differences between the Yolov8 and Yolov9 variants, including the use of disparate convolution kernels that modulate both representational capacity and parameter overhead [45].

Table 5. Yolov8 and Yolov9 models' performance for each location and corresponding mAP50. Bold data shows the best-performing model for each dataset. Five Yolov9 variants with the suffix _CWi (i=1,2,3,4,5) were customized models. The mAP50 decrease percentage indicates the relative performance drop of the YOLOv9_cw5 lightweight model compared to the top-performing model.

Model	Combined	Fargo 2022	Casselton 2021	Carrington 2022	Carrington 2021
Yolov8l	0.802	0.714	0.836	0.535	0.813
Yolov8m	0.784	0.696	0.831	0.543	0.821
Yolov8n	0.806	0.709	0.847	0.419	0.822
Yolov8s	0.806	0.611	0.833	0.419	0.81
Yolov8x	0.809	0.683	0.838	0.546	0.802
Yolov8 × 6	0.807	0.701	0.833	0.412	0.815
Yolov9	0.837	0.724	**0.862**	0.613	0.847
Yolov9c	**0.839**	**0.757**	0.851	0.616	**0.854**
Yolov9e	0.837	0.754	0.855	**0.627**	**0.854**
Yolov9_cw1*	0.834	0.716	0.858	0.622	0.846
Yolov9_cw2*	0.828	0.721	0.846	0.57	0.834
Yolov9_cw3*	0.823	0.678	0.854	0.565	0.818
Yolov9_cw4*	0.816	0.689	0.844	0.527	0.831
Yolov9_cw5*	0.821	0.699	0.85	0.424	0.835
mAP50 decrease %	2.15	7.66	1.39	32.38	2.22

Through deliberate customization, the Yolov9 architecture was slimmed down to a smaller memory profile and fewer trainable parameters, giving rise to five variants: Yolov9_cw1, Yolov9_cw2, Yolov9_cw3, Yolov9_cw4, and Yolov9_cw5. Compressing the model in this manner holds promise for trimming inference latency [27, 29]. Despite aggressive parameter and size cutbacks, our tailored Yolov9 derivatives maintained respectable accuracy. For example, our Yolov9_CW5 model has the second-lowest parameter count overall (after Yolov8n) yet beats Yolov8n on all datasets except Fargo 2022. Measuring the best-performing model from **Table 5** against our lightweight Yolov9_CW5 reveals a more pronounced mAP50 drop for the Carrington 2022 and Fargo 2022 datasets, while the performance gap for the remaining datasets remained modest—under 2.22%. The sensitivity of the Yolo model's performance to the underlying dataset underscores the centrality of data in constructing effective weed-detection systems. This viewpoint is buttressed by the No Free Lunch Theorems (NFL), which stipulate that no single optimization algorithm universally outperforms

© 2024 The Author(s). | Published by Agro Intelligence Research Press

others across all problem landscapes [46]. The NFL theorems thus advocate a data-centric mindset for crafting robust deep learning solutions.

Table 6 presents our lightweight model's performance on the combined dataset, side-by-side with results from prior published studies. Crucially, none of the earlier efforts listed in **Table 6** targeted crop species identification, whereas our detector can distinguish between eight crop and five weed species. This highlights the expanded scope and heightened complexity of our data, imposing a stiffer challenge on Yolo models to finely resolve diverse weed and crop classes. Aside from the works by [29, 47], every other study featured fewer than 5 classes, suggesting their datasets were less multifaceted than the one used here. Despite the elevated complexity, our custom lightweight model achieved mAP50 above 82%, signaling strong potential for real-time application across varied cropping systems. Beyond raw data intricacy, model complexity (measured via GFLOPs) is equally decisive for real-time weed detection. As a comparison point, whereas the Yolov4 model featured in [29] **(Table 6)** returned a marginally higher mAP@0.5, our detector offered a clear computational advantage, requiring just 26.5 GFLOPs versus Yolov4's roughly 148 GFLOPs. Our prior investigation [48] using the same dataset achieved a peak mAP@0.5 of 82% with Yolov4 on Casselton 2021 data, but at substantially higher computational cost—performance virtually equivalent to our Yolov9-cw1 lightweight variant. This exemplifies how newer Yolo generations can achieve a superior balance between accuracy and efficiency. Even with the parameter cutbacks in Yolov9-cw1, the model exhibited steady performance, indicating retained feature extraction quality. Such sturdiness can be attributed to regularization effects from using fewer filters, bottleneck phenomena, and the inherent robustness of the YOLOv9 design [38, 49-51].

Table 6. Comparison of the performance of customized YOLOv9 models with a previous study.

Research study	Models	Performance
Our Study	Yolov9_cw1 & Yolov9_cw5	The lightweight Yolov9_cw5 achieved an mAP50 of 82.1%, whereas the standard Yolov9_cw1 achieved 83.4%.
[47]	Yolov3, Yolov5, and Faster R-CNN	Yolov5 achieved the best average precision, reaching 92.4%.
[52]	Trained the CenterNet model	The model achieved a precision of 95.6%, a recall of 95%, and an F1-score of 0.953.
[29]	Yolov3, Yolov4, Scaled-Yolov4, YoloR, Yolov5, Yolov6, and Yolov7	Yolov4 achieved the best performance, with an mAP@0.5 of 95.03%.
[53]	Yolov5, Yolov6, Yolov7, and Yolov8	Yolov8l recorded the highest accuracy, with a precision of 0.9476 and mAP@0.5 of 0.9795.
[54]	Yolov3, Yolov4, and Yolov5	Yolov5s achieved the best results, with an F1-score of 75.3% and an accuracy of 77%.
[9]	Yolov5, RetinaNet, EfficientDet, Fast RCNN, and Faster RCNN	RetinaNet produced the best detection performance with an mAP@0.5 of 79.98 %.
[55]	MobileNetV2, ResNet50, and custom CNN models including 4-layer, 5-layer, and 8-layer CNN networks	The custom 5-layer CNN network achieved the highest detection accuracy at 97.7 %.
[56]	VGGNet, DetectNet, and GoogLeNet	DetectNet outperformed the other approaches, achieving an F1-score greater than 0.99.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

[28]	Custom Yolo-Spot_S, Yolo-Spot_M, and Yolo-Spot_L based on the Yolov7-tiny architecture	Yolo-Spot_M achieved the best precision of 87.4%, improving performance by 1.3% compared with the Yolov7-Base model.

The patterns observed across all five datasets matched our expectations. As shown in **Figure 7**, when crop foliage intertwines, placing predicted bounding boxes at the exact coordinates we originally labeled becomes problematic. This scenario would likely play out differently were weed and crop plants fully separated. In **Figure 7**, the broad-stroke predictions were on target, yet the spatial assignment of soybean plants departed from the manual annotations. This means that when one seeks outputs with tight bounding-box correspondence (high IoU), the odds of overlooking an otherwise correct soybean detection increase. Consequently, within weed and crop object detection tasks where canopies overlap, performance metrics can dip even though the detections themselves are sound. The model behavior documented here exposes the boundaries of a model-centric development mindset, highlighting the imperative for a data-centric approach that leverages extensive, high-fidelity datasets and enforces accurate spatial labeling during training [14].

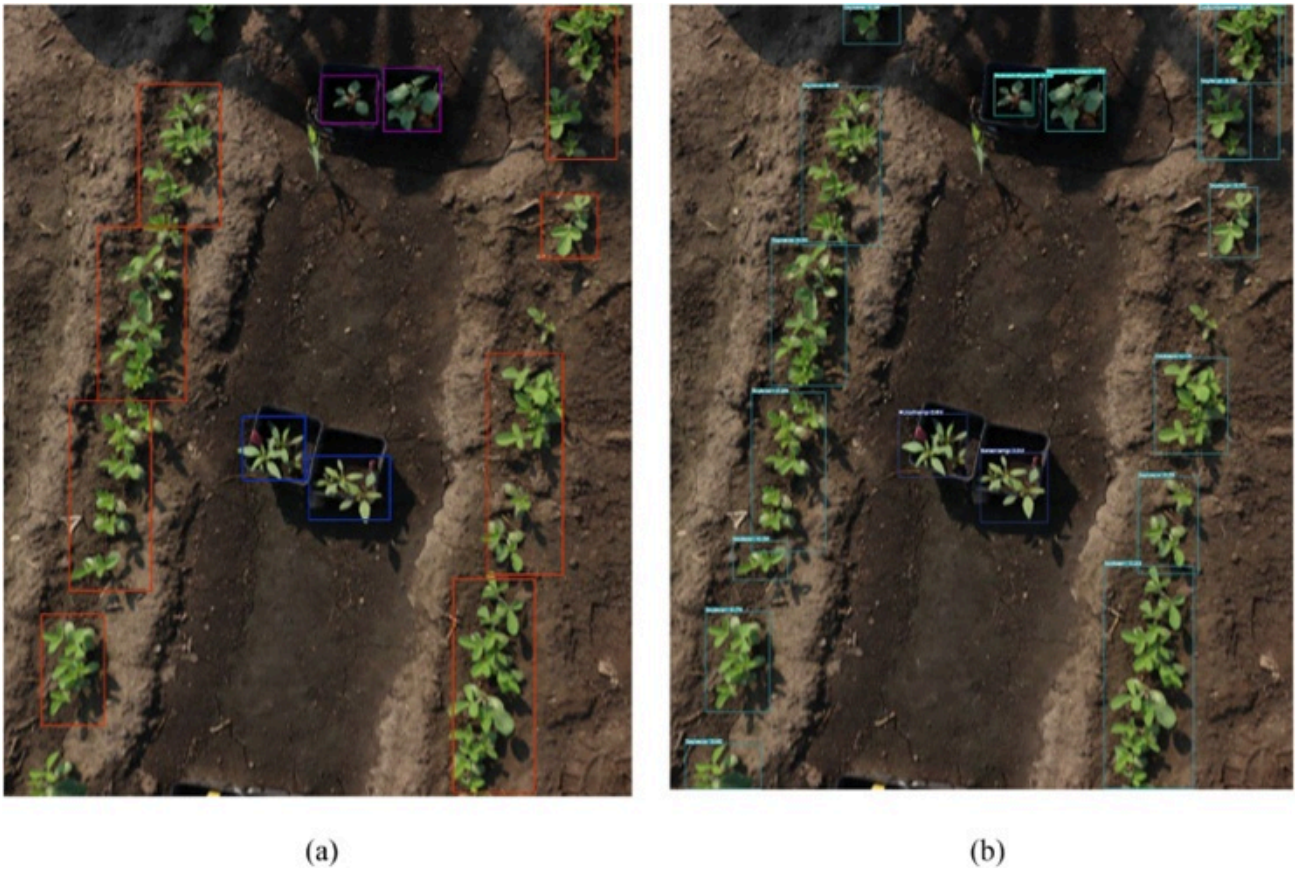


Figure 7. Difference between the labeled image (ground truth) and predicted image from the Yolov9_CW5 trained model: (a) Labeled image screenshot from LblImg software; orange bounding box is for Soybean, blue bounding box is for Waterhemp, and pink bounding box is for Redroot Pigweed; (b) Yolov9_CW5 prediction results from testing image: Soybean bounding box is predicted as soybean, and the pot between soybean rows is detected as redroot and waterhemp accurately.

Yolo model comparison based on crop and weed species

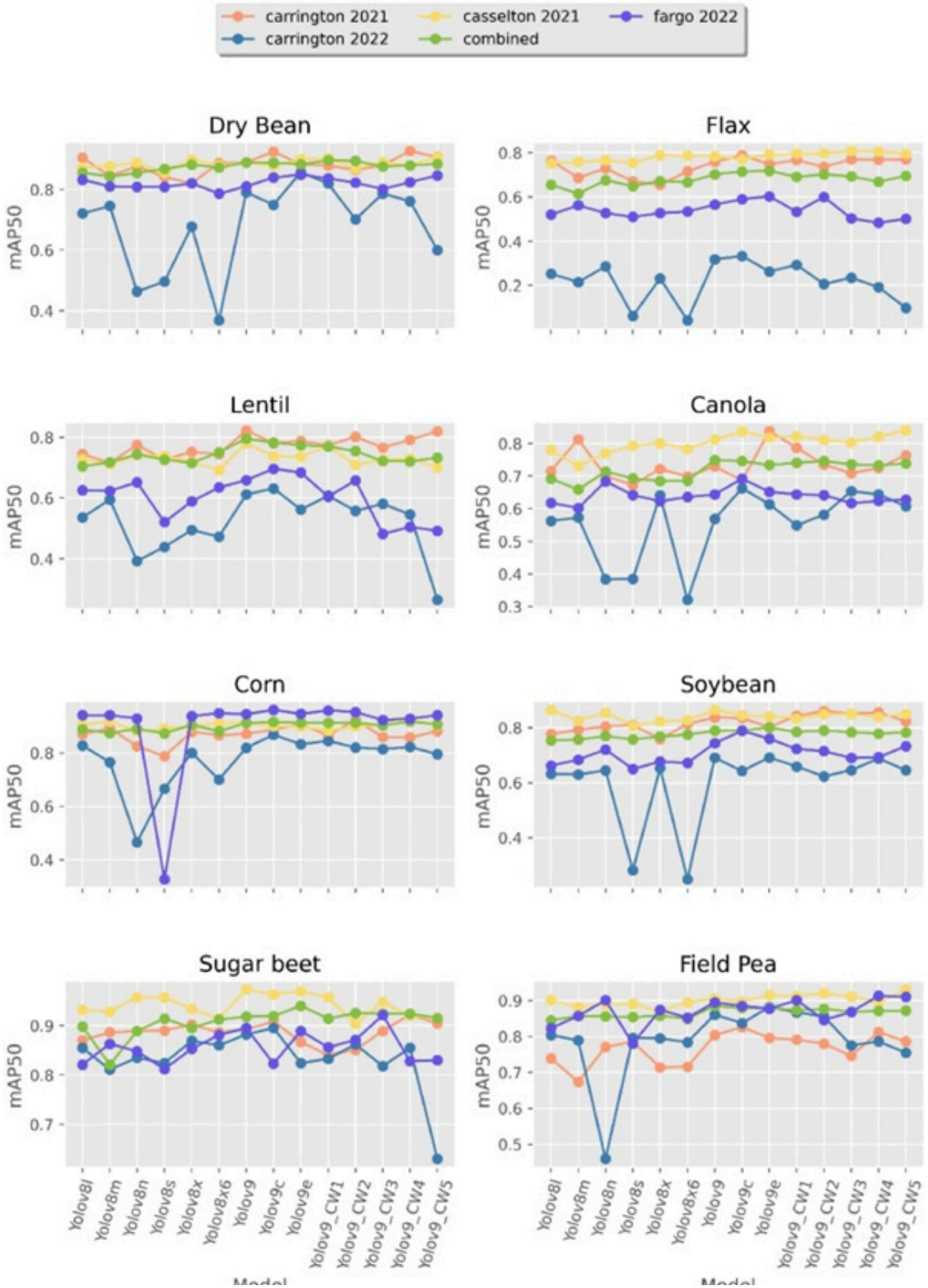
The trustworthiness of multi-class weed detection powered by deep learning can vary across target organisms, as crops and weeds often yield different accuracy profiles. Overlaps and divergences in morphological signatures shared by weeds

© 2024 The Author(s). | Published by Agro Intelligence Research Press

and crops intensify this challenge. A case in point is separating species with near-identical visual traits—waterhemp, Palmer amaranth, and redroot pigweed [30]. Farm operators generally focus weed suppression measures on a narrow roster of crop and weed types. Hence, tailoring models expressly for those species and balancing architectural decisions (model-centric) against data soundness (data-centric) could substantially support well-informed tactical choices.

The side-by-side assessment of the various Yolo configurations deployed here for discriminating crop and weed species is presented in **Figures 8 and 9**. Test-set evaluation reveals that the strongest figures came from Carrington 2021, Casselton 2021, and the merged corpus, where mAP50 exceeded 0.7 (70%) across virtually all crop and weed classes. In sharp contrast, every Yolo model struggled on the Carrington 2022 and Fargo 2022 collections, with mAP50 sinking below 0.7 (70%) for all categories except corn, sugar beet, field pea, dry bean, and horseweed. The data also showed that waterhemp identification fell below 0.6 (60%), while redroot pigweed identification even dropped below 0.5 (50%) when assessed on the Carrington 2022 and Fargo 2022 datasets.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

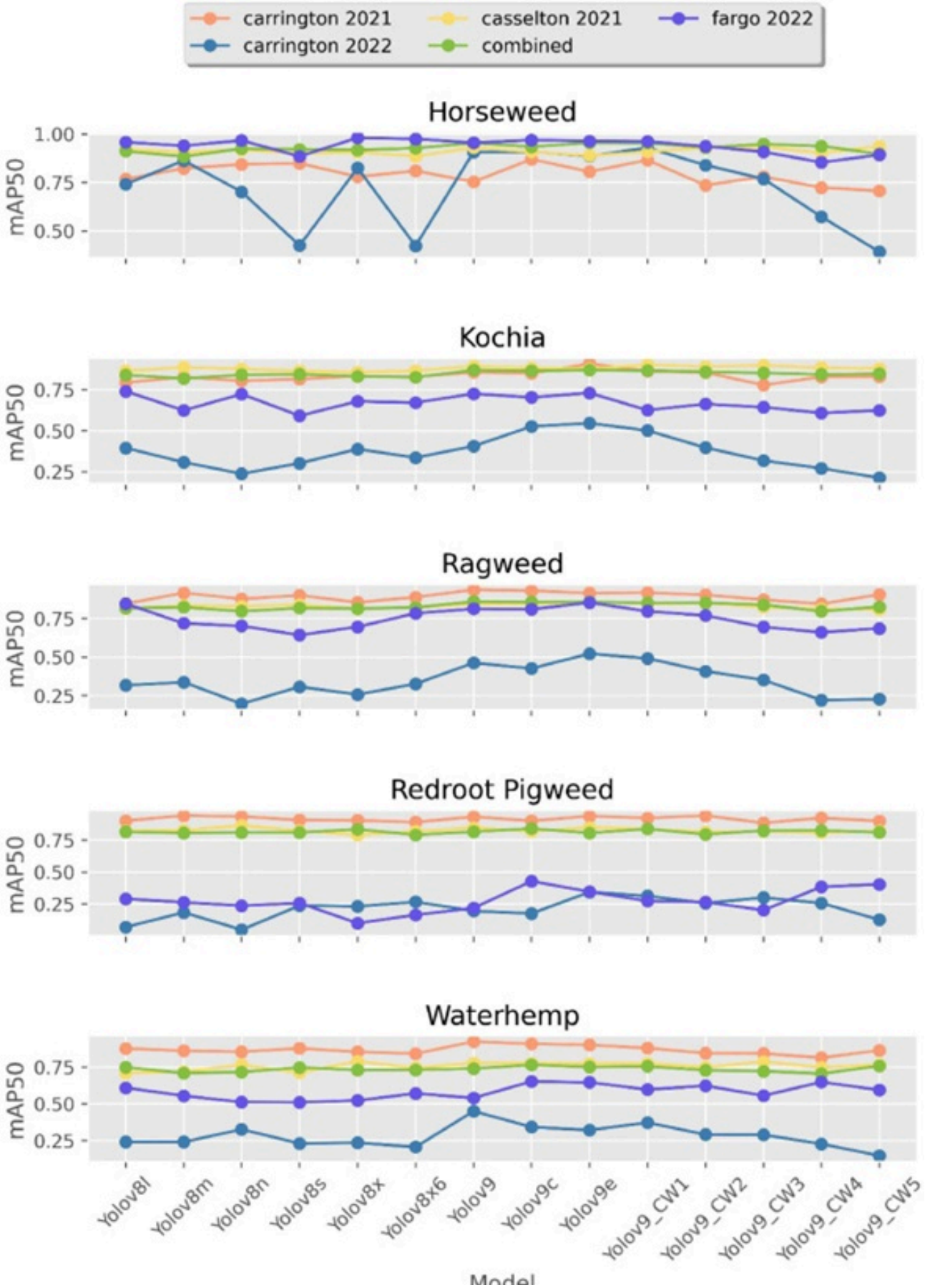


© 2024 The Author(s). | Published by Agro Intelligence Research Press



Figure 8. The performance of Yolov8 and Yolov9 models on five datasets representing eight crop production systems.

© 2024 The Author(s). | Published by Agro Intelligence Research Press



© 2024 The Author(s). | Published by Agro Intelligence Research Press

Figure 9. The performance of Yolov8 and Yolov9 models on five datasets representing five weed species.

Such weak results stem from a constrained dataset volume and a narrow visual variety, which cannot replicate real-world heterogeneity. To address this shortfall and introduce greater variability, we fused imagery from all four field sites into what we call the combined dataset. Training on this amalgamated collection improved all Yolo models to mAP50 values of 0.7 (70%) or higher across all crop and weed species, with Flax and Canola as the sole outliers. For those two crop types, select Yolo variants hovered between 0.6 (60%) and 0.7 (70%). Across the combined data, sugar beet and horseweed achieved peak accuracy, with mAP50 around 0.9 (90%) across all Yolo variants. These lofty detection rates for sugar beet and horseweed presumably owe to their uniquely identifiable shape profiles relative to the other species examined. At the other end, Flax consistently produced the lowest scores, with mAP50 remaining below 0.7 (70%) across all tested models, a result that may reflect compromised flax data quality due to environmental instability.

Results from the combined dataset suggest tightly clustered performance across all models for crop and weed classification tasks. Among the architectures examined, YOLOv9e achieved the highest accuracy for sugar beet, horseweed, kochia, flax, and soybean across the 13 species studied. The purpose-built variants Yolov9_CW1 and Yolov9_CW4 delivered the best outcomes for dry bean and corn, respectively. Meanwhile, the custom models Yolov9_CW1 and Yolov9_CW2 placed second for horseweed, redroot pigweed, sugar beet, canola, and soybean, as illustrated in **Figures 8 and 9**. These findings reinforce the ability of tailored YOLOv9 configurations to deliver reliable weed detection in specific crop contexts.

Table 7 places the current study alongside recent investigations focused on classifying and localizing a range of crop and weed species. The comparison identifies the top-performing model and dataset source for each crop and weed class featured here. The accuracy levels documented in this work are consistent with earlier published findings across all species. Flax, lentil, and dry bean had no prior research references for direct comparison. **Table 7** highlights the edge our customized model holds for field pea, canola, dry bean, flax, and redroot pigweed. These findings imply that assembling an ensemble of deep learning models could maximize detection performance tailored to specific crop or weed species. The strongest-performing DL model for a given weed–crop scenario could then be loaded onto machine-vision-driven precision spot sprayers to carry out site-specific weed management.

Table 7. Comparison of crop and weed species in our study with the previous studies. Our study shows the highest mAP50 value for weed and crop species across five datasets and 14 models. N/A signifies no studies found for weed detection and localization.

Crop/Weed	Our study	Previous study
Sugar beet	YoloV9e achieved an mAP50 of 96.9% on the Casselton2021 dataset.	Improved region-based fully convolutional network (R-FCN): AP reaching 84.8% [57]; Yolov3: mAP50 up to 91.5% [26]
Soybean	YoloV9 achieved an mAP50 of 86.5% on the Casselton2021 dataset.	Yolov5 reported an AP of 75.6% [10]
Corn	YoloV9c achieved an mAP50 of 96.2% on the Fargo2022 dataset.	Yolov4 achieved an AP of 86.67% [24]
Field Pea	YoloV9_CW5 achieved an mAP50 of 93.1% on the Casselton2021 dataset.	Yolov3 achieved an accuracy as high as 90.4% [21]
Canola	YoloV9_CW5 achieved an mAP50 of 84% on the Casselton2021 dataset.	DeepVeg DL model obtained an IOU score of 0.74 [58]
Lentil	YoloV9 achieved an mAP50 of 82.3% on the	N/A

© 2024 The Author(s). | Published by Agro Intelligence Research Press

	Carrington2021 dataset.	
Dry bean	YoloV9_CW4 achieved an mAP50 of 92.8% on the Carrington2021 dataset.	N/A
Flax	YoloV9_CW3 obtained an mAP50 of 80.8% on the Casselton2021 dataset.	N/A
Redroot Pigweed	YoloV9_CW2 and Yolo8m both achieved an mAP50 of 94.2% on the Carving2021 dataset.	Yolov3: AP of 89.89% [22]; Yolov7-Base: 100% accuracy [28]
Kochia	YoloV9e achieved an mAP50 of 90.8% on the Carrington2021 dataset.	ResNet50-SegNet: detection accuracy up to 99% for kochia patches [59]; Yolo-Spot_M: 89% accuracy [28]
Horseweed	YoloV8x achieved an mAP50 of 98.1% on the Fargo2022 dataset.	Yolov5: classification accuracy of 0.85 [48]; Yolo-Spot_M: 72% accuracy [28]
Water hemp	YoloV9 achieved an mAP50 of 92.3% on the Carrington2021 dataset.	ResNeXt101 achieved an accuracy of 100% [27]
Ragweed	YoloV9 achieved an mAP50 of 93.7% on the Carrington2021 dataset.	Yolov3: AP of 57.8% [22]; Yolo-Spot_M: 81% accuracy [28]; ResNeXt101: accuracy of 100% [27]

The performance divergence across models trained on the five separate datasets was expected, stemming from differences in sample counts, crop developmental stages, geographic site, and time of capture for the four source datasets [11]. As one example, weeds introduced in 2022 at Fargo and Casselton were notably taller than those transplanted in 2021 at Casselton and Carrington. Models trained on the amalgamated dataset likewise performed differently because that compilation pools data from every site, thereby injecting inherent variability. It is practically impossible to orchestrate field trials with identical environmental and soil conditions across multiple locations and years [60]. Climatic and edaphic variables can exert disparate influences on the germination, growth, and spatial patterning of weed and crop plants [60]. Over the course of this study, we similarly faced obstacles in collecting uniform imagery due to weather events, including rainstorms and dry spells, as well as inconsistent germination and growth rates across species; notably, certain crop seeds failed to thrive in the field.

Weed and crop detection model benefits, limitations, and future recommendations

The integration of deep learning (DL) object detection models into agriculture has grown steadily throughout the past decade. Current research trajectories indicate that most effort has been channeled into weed identification, with comparatively less focus on crop recognition. The present study contributes by tackling multi-crop species classification alongside multi-class weed detection. By accurately pinpointing crop species, more targeted weed management protocols can be implemented, reducing the likelihood of crop damage and supporting productivity gains.

The weed and crop detection models trained on Yolov8, Yolov9, and customized Yolov9 deep learning object detection variants have considerable potential to fundamentally alter weed control practices that currently rely on broadcast herbicide application. In this work, 14 localization-capable weed-detection models were trained on five datasets comprising five weed and eight crop species relevant to the North Dakota region of the United States. These detection models open avenues for diverse precision agriculture uses, including weed abatement, crop biomass estimation, and enabling autonomous vehicles to make informed chemical application decisions in the field. The encouraging performance demonstrated here sets the stage for creating hardware and software solutions aligned with grower requirements. For example, if a producer cultivates any of the eight crop species studied and wishes to manage weeds outside the five species investigated, the models could delineate weedy zones through straightforward image processing that subtracts crop areas.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Furthermore, if a grower seeks to focus on a particular weed, such as waterhemp, in a cropping system beyond the present study, herbicide application could be directed based on waterhemp detections generated by these models. Thus, the models offer considerable versatility to suit different on-farm weed control scenarios. Through this, both agricultural producers and the environment can benefit from site-specific precision weed management strategies that reduce herbicide consumption [61].

Practices such as model retraining on freshly collected data and iterative fine-tuning should be upheld to maintain a resilient artificial intelligence (AI) system across any domain, including agriculture. Accordingly, future work can incorporate additional data to address imbalances stemming from the inherently stochastic character of agricultural data. The models developed here carry certain constraints; foremost among them is their subdued performance on the Flax class, which could be strengthened in subsequent work through expanded data collection. Another constraint is the lack of coverage for weed species beyond the five considered, should a producer wish to zero in on a single species not represented. Consequently, the model could be retrained using images of additional weed species. The suite of constructed models could serve as interchangeable modules in weed-control robotic systems after fine-tuning and retraining, eliminating much of the overhead traditionally required for non-AI computer vision tasks. Lastly, the deep learning backbones employed in this study could be supplanted by more sophisticated, bespoke architectures to achieve still higher performance.

Conclusion

In the present study, site-specific weed detection models built upon Yolo object detectors were developed for five weed species (Horseweed, Kochia, Redroot Pigweed, Ragweed, and Waterhemp) across eight crop production systems (Dry Bean, Canola, Corn, Field Pea, Flax, Lentil, Soybean, and Sugar beet) commonly encountered in North Dakota, USA. The real-field image corpus underlying this work comprises data from four locations, gathered during the 2021 and 2022 seasons. These data served to train 14 Yolo object detection models, five of which were purpose-built for crop and weed classification. The investigation documented performance discrepancies attributable to variations in dataset size, site, and plant dimensions. Test results on the final merged dataset produced mean average precision (mAP50) values ranging from 0.784 (78.4%) for Yolov8m to 0.839 (83.9%) for Yolov9c. Moreover, the custom-built models achieved the highest detection success for corn and dry bean, with mAP50 values exceeding 0.89 (89%) in both cases. The results revealed wide performance swings in multi-species weed and crop detection from single images, driven by environmental factors, data volume, and intrinsic data properties. The custom lightweight model (Yolov9_CW5) outperformed every Yolov8 variant and approached the accuracy levels of the standard Yolov9 models, registering an mAP50 of 0.821 (82.1%) on the combined dataset. This outcome signals the promise of custom, lightweight models for embedding in edge devices for real-time functions, such as intelligent variable-rate herbicide spraying technologies, enabling sustainable weed control with minimal chemical use. Future research directions include deploying lightweight Yolov9 models within a vision-based weed-detection platform to dispense herbicides based on weed presence and evaluating the accuracy–speed trade-off under real field operating conditions.

Acknowledgements

This material is based on work partially supported by the U.S. Department of Agriculture under agreement number 58-6064-8-023. Any opinions, findings, conclusions, or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the view of the U.S. Department of Agriculture. This work is/was supported by the USDA National Institute of Food and Agriculture, Hatch project number ND01487.

Conflict of interest

The authors whose names are listed immediately below certify that they have NO affiliations with or involvement in any organization or entity with any financial interest (such as honoraria; educational grants; participation in speakers' bureaus;

© 2024 The Author(s). | Published by Agro Intelligence Research Press

membership, employment, consultancies, stock ownership, or other equity interest; and expert testimony or patent-licensing arrangements), or non-financial interest (such as personal or professional relationships, affiliations, knowledge or beliefs) in the subject matter or materials discussed in this manuscript.

Financial support

None

Ethics statement

None

Received: 03 Feb 2023 Revised: 10 Apr 2023 Accepted: 26 Apr 2023
Published online: 25 May 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

1. Ayoub Shaikh T, Rasool T, Rasheed Lone F. Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming. Comput Electron Agric. 2022;198:107119.

2. Yu Y, Zhang K, Yang L, Zhang D. Fruit detection for strawberry harvesting robot in non-structural environment based on Mask-RCNN. Comput Electron Agric. 2019;163:104846.

3. Kim EungChan, Hong Suk-Ju, Kim Sang-Yeon, Lee Chang-Hyup, Kim Sungjay, Kim Hyuck-Joo, et al. CNN-based object detection and growth estimation of plum fruit (Prunus mume) using RGB and depth imaging techniques. Sci Rep. 2022;12(1):1–16.
https://doi.org/10.1038/s41598-022-25260-9..

4. Ahmad I, Yang Y, Yue Y, Ye C, Hassan M, Cheng X, et al. Deep learning based detector YOLOv5 for identifying insect pests. Appl Sci. 2022;12(19):10167.
https://doi.org/10.3390/APP121910167..

5. Wang H, Shang S, Wang D, He X, Feng K, Zhu H. Plant disease detection and classification method based on the optimized lightweight YOLOv5 model. Agriculture. 2022;12(7):931.
https://doi.org/10.3390/AGRICULTURE12070931..

6. Ribeiro A, Bengochea-Guevara JM, Montes H, Sanchez PR, Zhang H. Evaluation of a CNN-based modular precision sprayer in broadcast-seeded field. Sensors. 2022;22(24):9723.
https://doi.org/10.3390/S22249723..

7. Wang T, Chen B, Zhang Z, Li H, Zhang M. Applications of machine vision in agricultural robot navigation: A review. Comput Electron Agric. 2022;198:107085.

8. Gerhards R, Andújar Sanchez D, Hamouz P, Peteinatos GG, Christensen S, Fernandez-Quintanilla C. Advances in site-specific weed management in agriculture—a review. Weed Res. 2022;62(2):123–33. https://doi.org/10.1111/WRE.12526..

9. Rahman A, Lu Y, Wang H. Performance evaluation of deep learning object detectors for weed detection for cotton. Smart Agric Technol. 2023;3:100126. https://doi.org/10.1016/J.ATECH.2022.100126..

10. Leon RG, Dille JA, Barnhart IH, Lancaster S, Goodin D, Spotanski J. Use of open-source object detection algorithms to detect Palmer amaranth (Amaranthus palmeri) in soybean. Weed Sci. 2022;70(6):648–62. https://doi.org/10.1017/WSC.2022.53..

11. Liu L, Ouyang W, Wang X, Fieguth P, Chen J, Liu X, et al. Deep learning for generic object detection: a survey. Int J Comput Vis. 2020;128(2):261–318. https://doi.org/10.1007/S11263-019-01247-4/FIGURES/21..

12. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit. 2016;2016-Decem:779–88. https://doi.org/10.1109/CVPR.2016.91..

13. Bochkovskiy A, Wang CY, Liao HYM. YOLOv4: optimal speed and accuracy of object detection. 2020. https://doi.org/10.48550/arxiv.2004.10934..

14. Diwan T, Anirudh G, Tembhurne JV. Object detection using YOLO: challenges, architectural successors, datasets and applications. Multimed Tools Appl. 2022:1–33. https://doi.org/10.1007/S11042-022-13644-Y/TABLES/7..

15. Wang Q, Cheng M, Huang S, Cai Z, Zhang J, Yuan H. A deep learning approach incorporating YOLO v5 and attention mechanisms for field real-time detection of the invasive weed Solanum rostratum Dunal seedlings. Comput Electron Agric. 2022;199:107194. https://doi.org/10.1016/J.COMPAG.2022.107194..

16. Chen J, Wang H, Zhang H, Luo T, Wei D, Long T, et al. Weed detection in sesame fields using a YOLO model with an enhanced attention mechanism and feature fusion. Comput Electron Agric. 2022;202:107412.

17. Zhao J, Tian G, Qiu C, Gu B, Zheng K, Liu Q. Weed detection in potato fields based on improved YOLOv4: optimal speed and accuracy of weed detection in potato fields. Electronics. 2022;11(22):3709. https://doi.org/10.3390/ELECTRONICS11223709.

18. Zhu H, Zhang Y, Mu D, Bai L, Zhuang H, Li H. YOLOX-based blue laser weeding robot in corn field. Front Plant Sci. 2022;13:3661.

19. Yang J, Wang Y, Chen Y, Yu J. Detection of weeds growing in alfalfa using convolutional neural networks. Agronomy. 2022;12(6):1459. https://doi.org/10.3390/AGRONOMY12061459.

20. Asad MH, Bais A. Weed detection in canola fields using maximum likelihood classification and deep convolutional neural network. Inf Process Agric. 2020;7(4):535–45. https://doi.org/10.1016/J.INPA.2019.12.002.

21. Khan S, Tufail M, Khan MT, Khan ZA, Anwar S. Deep learning-based identification system of weeds and crops in strawberry and pea fields for a precision agriculture sprayer. Precis Agric. 2021.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

https://doi.org/10.1007/s11119-021-09808-9.

22. Ahmad A, Saraswat D, Aggarwal V, Etienne A, Hancock B. Performance of deep learning models for classifying and detecting common weeds in corn and soybean production systems. Comput Electron Agric. 2021;184:106081. https://doi.org/10.1016/J.COMPAG.2021.106081.

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

24. Pei H, Sun Y, Huang H, Zhang W, Sheng J, Zhang Z. Weed detection in maize fields by UAV images based on crop row preprocessing and improved YOLOv4. Agriculture. 2022;12(7):975. https://doi.org/10.3390/AGRICULTURE12070975.

25. Peng H, Li Z, Zhou Z, Shao Y. Weed detection in paddy field using an improved RetinaNet network. Comput Electron Agric. 2022;199:107179. https://doi.org/10.1016/J.COMPAG.2022.107179.

26. Gao J, French AP, Pound MP, He Y, Pridmore TP, Pieters JG. Deep convolutional neural networks for image-based Convolvulus sepium detection in sugar beet fields. Plant Methods. 2020;16(1):29. https://doi.org/10.1186/s13007-020-00570-z.

27. Chen D, Lu Y, Li Z, Young S. Performance evaluation of deep transfer learning on multi-class identification of common weed species in cotton production systems. Comput Electron Agric. 2022;198:107091. https://doi.org/10.1016/J.COMPAG.2022.107091.

28. Rai N, Zhang Y, Villamil M, Howatt K, Ostlie M, Sun X. Agricultural weed identification in images and videos by integrating optimized deep learning architecture on an edge computing technology. Comput Electron Agric. 2024;216:108442. https://doi.org/10.1016/J.COMPAG.2023.108442.

29. Dang F, Chen D, Lu Y, Li Z. YOLOWeeds: A novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. Comput Electron Agric. 2023;205:107655. https://doi.org/10.1016/J.COMPAG.2023.107655.

30. GC S, Zhang Y, Howatt K, Schumacher LG, Sun X. Multi-species weed and crop classification comparison using five different deep learning network architectures. J ASABE. 2024;67(2):43–55. https://doi.org/10.13031/ja.15590.

31. Lin T. “LabelImg. https://github.com/tzutalin/labelImg, 2015.

32. Montagnani C, Gentili R, Smith M, Guarino MF, Citterio S. The worldwide spread, success, and impact of ragweed (Ambrosia spp.). CRC Crit Rev Plant Sci. 2017;36(3):139–78. https://doi.org/10.1080/07352689.2017.1360112.

33. Schramski JA, Sprague CL, Patterson EL. Environmental cues affecting horseweed (Conyza canadensis) growth types and their sensitivity to glyphosate. Weed Sci. 2021;69(4):412–21. https://doi.org/10.1017/WSC.2021.27.

34. Schwinghamer TD, Van Acker RC. Emergence timing and persistence of Kochia (Kochia scoparia). Weed Sci. 2008;56(1):37–41. https://doi.org/10.1614/WS-07-098.1.

35. Ghorbani R, Seel W, Leifert C. Effects of environmental factors on germination and emergence of Amaranthus retroflexus. Weed Sci. 1999;47(5):505–10. https://doi.org/10.1017/S0043174500092183.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

36. Guo P, Al-Khatib K. Temperature effects on germination and growth of redroot pigweed (Amaranthus retroflexus), Palmer amaranth (A. palmeri), and common waterhemp (A. rudis). Weed Sci. 2003;51(6):869–75. https://doi.org/10.1614/P2002-127.

37. Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. 2023.

38. Wang CY, Yeh IH, Liao HYM. YOLOv9: learning what you want to learn using programmable gradient information. 2024.

39. Upadhyay A, G C S, Zhang Y, Koparan C, Sun X. Development and evaluation of a machine vision and deep learning-based smart sprayer system for site-specific weed management in row crops: an edge computing approach. J Agric Food Res. 2024;18:101331. https://doi.org/10.1016/J.JAFR.2024.101331.

40. Sohan M, Sai Ram T, Rami Reddy CV. A review on YOLOv8 and its advancements. In: Jacob IJ, Piramuthu S, Falkowski-Gilski P, editors. Data Intelligence and Cognitive Informatics. Singapore: Springer Nature Singapore; 2024. p. 529–45.

41. Powers DMW. Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. J Mach Learn Technol. 2011.

42. Blaschko MB, Lampert CH. Learning to localize objects with structured output regression. In: European Conference on Computer Vision. Springer; 2008. p. 2–15.

43. Ahmed M, Hashmi KA, Pagani A, Liwicki M, Stricker D, Afzal MZ. Survey and performance analysis of deep learning based object detection in challenging environments. Sensors (Basel). 2021;21(15). https://doi.org/10.3390/S21155116.

44. Talaei Khoei T, Ould Slimane H, Kaabouch N. Deep learning: systematic review, models, challenges, and research directions. Neural Comput Appl. 2023;35(31):23103–24.

45. Zhao X, Wang L, Zhang Y, Han X, Deveci M, Parmar M. A review of convolutional neural networks in computer vision. Artif Intell Rev. 2024;57(4):1–43.

46. Wolpert DH, Macready WG. No free lunch theorems for optimization. IEEE Trans Evol Comput. 1997;1(1):67–82. https://doi.org/10.1109/4235.585893.

47. Wang P, Tang Y, Luo F, Wang L, Li C, Niu Q, et al. Weed25: A deep learning dataset for weed identification. Front Plant Sci. 2022;13:1–14. https://doi.org/10.3389/fpls.2022.1053329.

48. Gc S. Integrating high-throughput phenotyping, robotic, and artificial intelligence technologies in weed species identification. North Dakota State University. 2023.

49. Li H, Samet H, Kadav A, Durdanovic I, Graf HP. Pruning filters for efficient ConvNets. 5th Int Conf Learn Represent ICLR 2017 - Conf Track Proc. 2016. https://arxiv.org/abs/1608.08710v3.

50. Liu Z, Li J, Shen Z, Huang G, Yan S, Zhang C. Learning efficient convolutional networks through network slimming. Proc IEEE Int Conf Comput Vis. 2017;2017-October:2755–63. https://doi.org/10.1109/ICCV.2017.298.

51. Luo JH, Wu J, Lin W. ThiNet:A filter level pruning method for deep neural network compression. Proc IEEE Int Conf Comput Vis. 2017. p. 5068–76. https://doi.org/10.1109/ICCV.2017.541.

52. Jin X, Sun Y, Che J, Bagavathiannan M, Yu J, Chen Y. A novel deep learning-based method for detection of weeds in vegetables. Pest Manag Sci. 2022;78(5):1861–9.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

https://doi.org/10.1002/PS.6804.

53. Sportelli M, Apolo-Apolo OE, Fontanelli M, Frasconi C, Raffaelli M, Peruzzi A, et al. Evaluation of YOLO object detectors for weed detection in different turfgrass scenarios. Appl Sci. 2023;13(14). https://doi.org/10.3390/app13148502.

54. Pérez-Porras FJ, Torres-Sánchez J, López-Granados F, Mesas-Carrascosa FJ. Early and on-ground image-based detection of poppy (Papaver rhoeas) in wheat using YOLO architectures. Weed Sci. 2023;71(1):50–8. https://doi.org/10.1017/wsc.2022.64.

55. Razfar N, True J, Bassiouny R, Venkatesh V, Kashef R. Weed detection in soybean crops using custom lightweight deep learning models. J Agric Food Res. 2022;8:100308. https://doi.org/10.1016/J.JAFR.2022.100308.

56. Yu J, Sharpe SM, Schumann AW, Boyd NS. Deep learning for image-based weed detection in turfgrass. Eur J Agron. 2019;104:78–84. https://doi.org/10.1016/J.EJA.2019.01.004.

57. Sun J, Yang K, He X, Luo Y, Wu X, Shen J. Beet seedling and weed recognition based on convolutional neural network and multi-modality images. Multimed Tools Appl. 2022;81(4):5239–58. https://doi.org/10.1007/s11042-021-11764-5.

58. Das M, Bais A. DeepVeg: Deep learning model for segmentation of weed, canola, and canola flea beetle damage. IEEE Access. 2021;9:119367–80. https://doi.org/10.1109/ACCESS.2021.3108003.

59. Asad MH, Saleem S, Bais A. Detection and mapping of Kochia plants and patches using high-resolution ground imagery and satellite data: application of machine learning. IEEE Access. 2023;11:92298–311. https://doi.org/10.1109/ACCESS.2023.3308909.

60. Sheng Y, Zhao S, Yang S. Weather shocks, adaptation and agricultural TFP: a cross-region comparison of Australian Broadacre farms. Energy Econ. 2021;101:105417. https://doi.org/10.1016/J.ENECO.2021.105417..

61. Thompson JF, Stafford JV, Miller PCH. Potential for automatic weed detection and selective herbicide application. Crop Prot. 1991;10(4):254–9. https://doi.org/10.1016/0261-2194(91)90002-9.

© 2024 The Author(s). | Published by Agro Intelligence Research Press