ORIGINAL RESEARCH **Open access**

Image-Based Detection of Wheat Field Weeds from Emergence to Maturity Using YOLOv5

Elena Petrova[1*], Ivan Georgiev[1], Nikolay Stoyanov[2], Petar Kolev[3]

Abstract

Managing weeds efficiently in wheat (*Triticum aestivum* L.) is fundamental for boosting production and securing food supplies in semi-arid environments. Deploying deep learning for weed recognition can support site-specific weed control, thereby enhancing wheat output, raising farmer profitability through precise herbicide usage, and lessening the ecological burden of chemical applications. Charlock mustard (*Sinapis arvensis* L.) [CM], creeping thistle (*Cirsium arvense* (L.) Scop) [CT], and forking larkspur (*Consolida regalis* Gray) [FL] rank among the most competitive weeds infesting wheat. Proper handling of these species is indispensable for sustaining wheat yields in semi-arid zones. An unmanned aerial vehicle collected imagery and video across 185 wheat fields in Tokat, Turkey, to facilitate weed identification. Image acquisition spanned five phenological phases for each weed: seedling, rosette, vegetative, flowering, and fruiting, totaling 15 distinct classification tasks. These developmental stages matter either because they represent critical windows for weed intervention or because they shed light on weed pressure thresholds, the risk of enlarging the weed seedbank, and the detection of herbicide-resistant biotypes. Annotations across the images produced a dataset of 145,792 labeled objects, partitioned into 80% for training plus validation and 20% held out for testing. This dataset was processed using the YOLOv5 (You Only Look Once) deep learning framework, which was the latest iteration at the time this work was performed. Every YOLOv5 neural network variant—nano, small, medium, large, and xlarge—underwent training, after which their Precision, Recall, F-1 Score, and AUC (Area under the Curve) metrics were compared. Across species, the YOLOv5s (small) network achieved the highest detection precision. In contrast, the nano architecture proved least accurate when detecting any single weed species. For CT, detection precision with the small network fluctuated between 0.87 at the fruit stage and 0.96 at both the rosette and vegetative stages. Regarding forking larkspur, the top-performing configuration was YOLOv5 small, where precision values stretched from 0.45 at the fruit stage to 0.93 at the vegetative stage. With Charlock mustard processed with YOLOv5 small, the best detection precision was observed at the seedling, rosette, and vegetative stages, each scoring 0.96, whereas the fruit stage recorded the lowest precision at 0.81. The trends observed in Recall, F-1 Score, and AUC closely mirrored those of Precision. These outcomes demonstrate that the YOLOv5 small neural network can detect all three weed species across all growth stages, with a precision of 0.86, except at the fruit stage. Upcoming work ought to (1) prioritize benchmarking the YOLOv5 small neural network against more recent YOLO versions and also (2) test its capability to detect weed species in wheat fields, thereby supporting enhanced weed management and informing future recommendations.

Keywords Deep learning, YOLOv5, Weed detection, Growing stages, Wheat

*Correspondence:
Elena Petrova
elena.petrova@au-plovdiv.bg
[1] Department of AI for Lavender and Rose Oil Production Optimization, Faculty of Agriculture, Agricultural University of Plovdiv, Plovdiv, Bulgaria
[2] Department of Smart Beekeeping and Pollination Mapping, Faculty of Agriculture, Sofia University, Sofia, Bulgaria
[3] Department of Precision Herbicide Application, Faculty of Agriculture, University of Ruse, Ruse, Bulgaria

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

Introduction

Achieving high crop yields and safeguarding global food supplies hinges on maximizing agricultural output while removing obstacles, such as weed competition, that depress productivity [1-3]. Food security on a global scale depends on a steady supply of several essential cash crops, among them Maize (*Zea mays* L.), rice (*Oryza sativa* L.), and wheat. Of these, wheat plays a prominent role in the diet across the Middle East, including Turkey. In Turkey, wheat is a cornerstone agricultural commodity that contributes significantly to export revenues and domestic self-reliance. Nevertheless, wheat quality and yield in Turkey are frequently undercut by diseases, insect pests, and weeds, all of which can dramatically shrink harvests. Raising aggregate crop output generally occurs either by expanding the cultivated area or by lifting yields per unit area. Because Turkey has no untapped farmland suitable for expansion, increasing crop yields is the only route to boosting production [4]. Yield improvement can be pursued through diverse avenues, including genetic gains, better soil stewardship, installation of more efficient irrigation infrastructure, and the application of crop protection products to restrain pests and diseases. Furthermore, the adoption of novel technologies and practices, such as precision farming, enables growers to fine-tune production by improving pest control, irrigation, and nutrient management [5].

Weeds compete with cash crops for essential resources, causing yield depression and financial setbacks for farmers. The severity of weed impact varies by crop species and geographic setting. Several investigations estimate the worldwide annual cost of weeds at 13.2% of potential output, equating to 75.6 billion USD. Country-specific losses amount to 2.7 billion USD in Australia, 5 billion USD in the United States, and 11 billion USD in India [6-9]. Under some circumstances, weed-induced damage can approach 100% when no control measures are applied, and even with intervention, technical shortcomings can still result in wheat yield losses reaching 20%. Deploying effective weed control programs is therefore vital for preserving crop productivity and curtailing economic losses [1, 10-12].

In Turkish wheat-producing areas, charlock mustard (*Sinapis arvensis*), creeping thistle (*Cirsium arvense*), and forking larkspur (*Consolida regalis*) are recognized as major weed adversaries responsible for substantial yield reductions [13-23]. Spotting weeds accurately is a high-priority activity for raising wheat yields and reining in these troublesome species. Weed detection lays the groundwork for crafting efficient management strategies. Having dependable data on weed density and spatial distribution helps forecast economic repercussions and supports the compilation of weed proliferation maps [24, 25]. Image processing coupled with artificial intelligence can streamline this workflow, making weed surveys simpler and less expensive than relying solely on expert visual inspections. A range of image processing and AI-centered methods have been documented for weed detection in the scientific record [26-33]. Such work typically involves capturing weed images with specialized instruments and later analyzing them on a computer. While manual scouting by trained personnel is an alternative, it carries inherent drawbacks. For example, the process can be slow, costly, and difficult to scale across large tracts. What is more, weed appearance shifts with environmental circumstances, further complicating manual identification [34].

Artificial Neural Networks (ANNs) are a widely adopted computer vision technology suited to object detection tasks. Through image-based training, ANNs can learn to identify patterns in data. El-faki *et al.* [26] exploited differences among color indices [red, green, blue (RGB), and gray scale] to separate weeds from soybean (*Glycine max* L.) and wheat. Weeds often display reddish roots, whereas wheat and soybeans appear green, enabling a straightforward color-based distinction. The weeds examined included Johnson grass (*Sorghum halepense*), redroot (*Amaranthus retroflexus*), and blackgrass (*Alopecurus myosuroides*) in soybean, and wild buckwheat (*Fallopia convolvulus*), cheatgrass (*Bromus tectorum*), and field bindweed (*Convolvulus arvensis*) in wheat. El-faki *et al.* [26] found that pairing color indices with discriminant analysis within AI classifiers enabled effective weed detection, though performance varied with weed type and developmental stage. They advised that as weeds age, the reddish root coloration fades, requiring stage-specific calibration of detection methods [26].

In a separate study, Bagheri *et al.* [35] applied machine learning tools, namely ANNs and canonical discriminant analysis (CDA), to construct a computer vision system for classifying weed seeds—*Sinapis arvensis*, *Chenopodium album*, *Alyssum hirstumu*, *Amaranthus blondes*, and *Amaranthus retroflexus*—based on physical attributes. The trained ANN

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

achieved approximately 69% accuracy in identifying *Sinapis arvensis* seeds. The authors also noted that seed shape characteristics proved especially valuable for weed categorization [35].

Support vector machines (SVMs) are another family of machine learning algorithms for pattern recognition and classification. In one experiment, Rumpf *et al.* [36] enlisted SVMs to tell apart two weed species (*Cirsium arvense* and *Galium aparine*) across three distinct phenological groupings (monocotyledons, dicotyledons, and barley). They used the RELIEF-F algorithm to extract features from plant imagery and the SVM-Weighting approach for classification. They reported classification accuracies of around 80% for both weed species [36].

Elsewhere, machine learning was paired with a robotic platform equipped with a hyperspectral camera to identify plants based on their spectral signatures. The system was taught using images of multiple weeds: *Alopecurus myosuroides*, *Cirsium arvense*, *Galium aparine*, *Lamium* sp., *Matricaria inodora*, *Sinapis arvensis*, *Stellaria media*, *Veronica persica*, and *Hordeum vulgare*. The top outcomes were delivered by integrating a mixture of Gaussians (OMG) with a self-organizing map (SOM) within an SVM. The research team reported successful identification rates of 94% for *Cirsium arvense* and 90% for *Sinapis arvensis* [37].

Published research indicates that spectral cameras can reliably distinguish weeds from crops by exploiting differences in radiation reflectance. In a Spanish study, deCastro *et al.* [38] employed hyperspectral and multispectral imagery to detect and classify winter wheat alongside weeds. Two supervised techniques—a vegetation index and the Maximum Likelihood Classifier (MLC)—were used for classification. Weeds were grouped broadly by leaf dimensions and form. By generating maps from these broad-scale classifications, the investigators were able to direct weed management efforts, trimming herbicide consumption by roughly 61% [38].

The popularity of UAVs in agriculture is growing, driven by technical advances, more affordable units as new manufacturers enter the scene, and straightforward operational controls. Rotary-wing UAVs, such as drones, exemplify this trend with their vertical takeoff capability and nimble maneuverability. In one investigation, researchers employed a drone to acquire overhead images. Their goal was to single out *Cirsium arvense* within a barley stand and produce a map for site-specific weed treatment. A range of flight altitudes (1-50 m) was tested, and imagery was analyzed to separate the weed from the crop. The strongest performance was observed in flights at 35 m [39].

Another research effort focused on mapping charlock mustard in alfalfa (*Medicago sativa* L.) fields. The drone deployed was outfitted with a Normalized Difference Vegetation Index (NDVI) sensor alongside a standard RGB camera. Imagery from both sensors underpinned computer vision processing and charlock mustard localization. The key differentiating trait was bloom color—yellow for charlock mustard versus green for alfalfa. The proposed dual-sensor method outperformed the RGB-only camera for producing weed-crop separation maps [40].

Deep learning stands as the most advanced form of AI and is exceptionally powerful for object detection. Contrary to classical methods, deep learning obviates the need for handcrafted feature extraction. Instead, it autonomously derives features and learns to use them for decision-making by referencing labeled objects in the image set. Özgür and Nar [41] used deep learning to recognize seeds of common crops (corn, wheat, and sugarcane [*Saccharum officinarum* L.]) and weeds (*Sisymbrium officinale*, Stellaria media, *Capsella bursa-pastoris*, *Galium aparine*, *Sinapis arvensis*, *Sedum* spp., and *Geranium molle*). They trained several deep learning architectures (LeNet5, Alexnet, VGG16, and MobileNet) on images acquired from a light-insulated photography platform and compared their results. MobilNet yielded the highest prediction accuracy, close to 86% [41].

A review of the literature reveals certain limitations inherent in earlier studies. For instance, investigations relying on image processing and statistical techniques have exploited color and shape differences to distinguish weeds from crops. However, since both weeds and crops pass through numerous phenological stages during their life cycle, certain shape and color features can overlap during particular phases. Consequently, such an approach may fail to deliver consistent accuracy across all phenological stages [42]. Additionally, considering the enormous diversity of weed taxa in nature, it is

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

critical to verify the reliability of weed detection and identification outcomes. Moreover, weed detection methods that rely on image processing and machine learning—such as ANNs and SVMs—require feature extraction to train the AI and achieve accurate classification [30, 43-45]. Yet agricultural images vary under the influence of abiotic factors such as weather, rainfall, and fluctuating daylight intensity, making it impossible to assign fixed bias values to human-extracted features in the neural network. Hence, even images shot from identical vantage points and altitudes of the same object can appear dissimilar, posing a genuine hurdle for these approaches.

Additionally, many weed detection studies have concentrated on only one or a subset of weed and crop phenological stages, even when employing deep learning. Some weeds, however, can look alike at certain stages (e.g., *Sinapis arvensis* and *Sisymbrium officinale*; *Consolida regalis* Gray and *Fumaria officinalis* L.), risking misclassification if examination is limited to those stages [33, 41, 46-49]. Conversely, some studies have used plant-weed images photographed under controlled conditions against plain backgrounds. In an authentic field setting, however, backgrounds are cluttered and can include assorted weeds, soil hues, stones, plant shadows, and so forth. It is therefore important to train an AI system using images that include both background clutter and target objects. An AI tool trained solely on controlled-background images is apt to lose accuracy when faced with a natural environment [32, 50, 51].

The purpose of the present study was to apply deep learning methods to detect common weeds—charlock mustard, creeping thistle, and forking larkspur—at different growth stages. These weeds are recognized for their significant toll on wheat production in Turkey and across the broader Middle East. To the best of our knowledge, this investigation constitutes the first and largest of its kind to detect principal wheat-field weeds from the seedling stage through fruiting using ANNs.

Materials and Methods

Site survey and data acquisition

This research focused on wheat production fields in Tokat, Turkey, during the 2019–2021 period, with the main goal of identifying three weed taxa—charlock mustard, creeping thistle, and forking larkspur—just before the 2019 harvest to establish a baseline reference. To fulfill this, a field reconnaissance was conducted, and a Magellan eXplorist 310 handheld GPS unit was used to construct a map showing the locations of the documented weeds. Seeds from these weeds were also harvested to propagate the species in a controlled setting, building a repository of image and video resources to draw upon if the weeds in the wheat fields failed to perform as hoped.

In parallel, video recordings were made with a DJI Mavic 2 Pro drone equipped with an RGB camera, delivering 3840×2160-pixel resolution. This camera setup captured the earliest phenological stage, marked by cotyledon leaves (seedling phase), at a flying height of 2 m. In this work, computer vision was used to recognize three weeds (charlock mustard, creeping thistle, and forking larkspur) across five phenological stages: seedling, rosette, vegetative, flowering, and fruiting.

Data pre-processing and dataset production method

Deep learning has seen a surge in scientific uptake thanks to detection results that outperform classical image processing pipelines [52-54]. Object detection grounded in deep learning requires hand-annotating the target entity in an image and subsequently training an ANN on these labeled instances. The richness of the training image set directly boosts predictive accuracy. Because weeds occur in dense stands across farmlands, amassing the sizeable datasets deep learning requires is relatively straightforward. This technique supports concurrent multi-object detection and assigns predicted items to predefined groupings. In the current study, the ambition was to recognize 15 unique weed categories, representing 5 phenological stages for each of the 3 weed species.

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

A balance in the number of objects per image was sought across categories; however, there were instances in which multiple objects were tagged within a single frame, as the photographs originated from actual wheat farms.

An Nvidia RTX 3090 GPU served as the core computing engine. Deep learning frameworks geared toward object detection can handle still images, video streams (individual frames), or perform live detection via sources such as webcams. Jocher *et al.* [55] highlight YOLO's edge over competing deep learning architectures, especially its frame-per-second (FPS) detection performance in video, which motivated our choice of YOLO for weed detection here. YOLO comes in several iterations, and the choice among them reflects user needs. YOLOv5 offers a spectrum of ANN architectures suited to the scale of the object under study [55]. The ANN variants included in YOLOv5 are nano, small, medium, large, and extra-large [44, 55].

The Confusion Matrix serves as an evaluation tool for assessing how well AI applications, including both machine learning and deep learning, perform on prediction and classification tasks [56]. It provides the data required to assess model correctness and compare different ANN designs using a shared dataset. Several mathematical expressions exist in the literature to measure AI application performance, among them **Eqs. 1-4** [44, 57-60].

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

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

$$F1\ score = \frac{2}{\frac{1}{Precision}+\frac{1}{Recall}} = \frac{2*(Precision*Recall)}{(Precision+Recall)} \quad (3)$$

$$AUC = 0.5\left(\frac{TP}{TP+FN} + \frac{TN}{TN+FP}\right) \quad (4)$$

In this study, the pictures and the accompanying .txt files comprising the training portion of the dataset were randomly split into two segments: a training batch and a validation batch, with 80% and 20%, respectively. Afterward, the trained deep learning framework was tested on images, video clips, and live video feeds withheld from training. Two distinct Python programs were employed to examine the trained deep learning frameworks:

• For video-based testing: "python detect.py –img 1024 – c:\yolov5\source cem.mp4 [the testing video path] –weights c:\yolov5\best.pt [any trained architecture path] –conf-thres 0.5″

• For real-time video testing: "python detect.py –weights c:\yolov5\best.pt [any trained architecture path] –source 0 [that source 0 uses the USB hub as an image-video source for any video-photo capture tools that can connect to the computer via USB cable].

Results and Discussion

The training regimen was initially set to 200 epochs, but it became evident that the model stopped improving after the 100th epoch. Accordingly, the epoch tally was reduced to 110 for each YOLOv5 variant, providing a 10% buffer beyond the noted stagnation point.

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

To enable AI applications (machine learning, deep learning) to forecast reliably, they need a “false class” against which to contrast genuine predictions. In this investigation, background scenes were included in the dataset to serve as false classes. Such images, gathered from real wheat-farming settings, contain various elements, including rocks, bare earth, wheat crops, other weed species, and farm machinery, forcing the system to acquire a nuanced understanding of what distinguishes the target weeds from everything else.

Based on the YOLOv5 nano model trained on our dataset, the CM2, CM3, and CT2 categories posted the highest true positive rates at 87%. At the other extreme, the FL5 and CM5 categories recorded the feeblest prediction outcomes of merely 8% and 18%, respectively.

Our YOLOv5s model, after training on the dataset, achieved the highest prediction accuracy for the CM, CM2, CM3, CT2, and CT3 categories (96%). The FL5 category, by comparison, produced the lowest prediction scores. Averaged across all categories, prediction accuracy topped 90%, signaling that the YOLOv5 model could serve as a trustworthy instrument for correctly flagging every weed category present in our dataset.

The YOLOv5 medium model schooled on our dataset attained its top prediction performance within the CT2 category at 93%. Trailing it were the CM2 and CM3 categories, both standing at a prediction rate of 91%. In contrast, the FL5 and CM5 categories displayed the poorest prediction rates, at 32% and 56%, respectively. Taken together, the average prediction rate across all categories remained above 80%, marking this model as a sound candidate for weed detection given our dataset.

When the trained YOLOv5 large model was applied to our dataset, it produced its best prediction scores for the CM2 and CT2 categories, both at 93%. Next in line were the CM1 and CM3 categories at 92% and 91%, respectively. The FL5 category, nonetheless, recorded the smallest prediction rate at a scant 35%.

The top prediction ratio for the YOLOv5 xlarge model educated on our dataset was observed in the CT2 category at 94%, shadowed by CM1, CM2, and CM3, which each logged 91% prediction accuracy. The weakest prediction results stemmed from the CM5 category at 16%.

The YOLOv5 models’ performance was evaluated using precision, recall, F1 score, and AUC. This evaluation relied on confusion matrix outputs for each model, and performance figures were tallied for each weed category and its associated phenological stages.

YOLOv5s (small) stood out as the most exact neural network for weed detection irrespective of species. On the flip side, nano proved the least exact neural network when detecting any single weed species. For CT, under the small neural network, detection precision spanned from 0.87 (fruit stage) up to 0.96 (rosette and vegetative stages). For forking larkspur, peak neural network performance was achieved with YOLOv5 small, with precision ranging from 0.45 (fruit stage) to 0.93 (vegetative stage). For Charlock mustard, handled by YOLOv5 small, the highest-precision detection occurred at the seedling, rosette, and vegetative stages (0.96 each), while the lowest-precision detection occurred at the fruit stage (0.81). The remaining performance yardsticks (Recall, F-1 Score, and AUC) were consistent with Precision. A positive linear linkage emerged between Precision and Recall (R^2 = 0.98), Precision and F-1 Score (R^2 = 0.99), and Precision and AUC (R^2 = 0.95). Our findings additionally revealed that by deploying the YOLOv5 small neural network, all three weed species could be spotted at each growth phase with a precision of 0.86, the fruit stage being the lone exception. Shifting to the medium, large, and xlarge architectures lowered weed detection precision by 8%, 7%, and 8%, respectively. This suggests that workable weed detection can be achieved with YOLOv5 versions beyond the small one, though at a roughly 8% drop in accuracy. Set against medium, large, and xlarge, choosing nano led to a 21% drop in weed species prediction precision, suggesting nano should be avoided for pinpointing the three selected weed species within wheat fields.

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

Deep Learning is a pivotal domain within Artificial Intelligence, one that has become fundamental for reliable weed identification, thereby enabling weed management tailored to specific locations [61]. In machine learning pipelines for weed recognition, characteristics such as leaf area, textural contrasts, and morphological traits are treated as explanatory variables to distinguish cash crops from weeds within an image [28, 62-66]. A recent survey of the literature by Rai *et al.* [61] and Murad *et al.* [67] firmly establishes that YOLO, an architecture purpose-built for image recognition and classification, surpasses alternative models for weed detection. Its enhanced performance makes it a go-to solution for precisely identifying and handling weeds across a variety of contexts. Although multiple YOLO generations can be applied to weed detection, this work focused on YOLOv5 and scrutinized five distinct weight configurations: nano, small, medium, large, and xlarge. Our outcomes were consistent with earlier published reports that point to YOLO as a powerful model for distinguishing weeds from cash crops. Our data confirmed that charlock mustard, creeping thistle, and forking larkspur were successfully detected in 185 wheat fields and in specimens cultivated in the laboratory. The element that sets our investigation apart from contemporary literature is the exhaustive assessment of weed detection across the entire growth phases, from seedling emergence through to fruiting. Unlike earlier works that confined themselves to a single phenological window or that sourced plant imagery under carefully regulated lighting and environmental conditions [33, 41, 46, 68-72], our research spanned a substantially wider range of developmental timepoints. This broader lens provides a more authentic and pragmatic gauge of weed-detection capability under shifting developmental conditions, thereby rendering our conclusions more directly transferable to real farming situations. Additionally, this tactic permitted us not only to flag weeds early in the cropping cycle—providing the necessary intelligence to evaluate weed pressure thresholds and execute timely interventions—but also to catch weeds at later phases that might otherwise shed seed and swell the soil seedbank. Such knowledge is indispensable for integrated weed management strategies within wheat-based cropping systems.

The findings of this study further revealed that YOLOv5s (small) was the most precise neural network for weed detection across species. This observation does not align with the published record, underscoring the importance of testing diverse weed species across dissimilar cash crop settings. As a case in point, Junior and Ulson [73] measured detection across multiple weed species and found that the YOLOv5s model performed worst, with only the categories small-leaved holly (0.59) and green amaranth (0.57) exceeding 0.5. They identified YOLOv5x as the most faithful network, achieving the strongest predictions for small-leaved holly (0.78), goose grass (0.84), and green amaranth (0.76) [73]. These patterns illustrate the uneven performance of YOLOv5 configurations when detecting specific weed classes, underscoring the need to evaluate models and weed species across diverse cash crop contexts. In our own work, nano emerged as the least precise neural network for detecting each weed species. For CT, with the small neural network, detection precision ran from 0.87 (fruit stage) up to 0.96 (rosette and vegetative stages). For forking larkspur, the top-performing neural network was YOLOv5 small, with precision ranging from 0.45 (fruit stage) to 0.93 (vegetative stage). When YOLOv5 small was applied to Charlock mustard, the highest precision was observed at the seedling, rosette, and vegetative stages (all with 0.96). In contrast, the least precise detection was recorded at the fruit stage (0.81). The documented ranges sit well with earlier reports in the literature suggesting that lightweight models tend to be more adept at weed detection. For example, Ruigrok *et al.* [74] leveraged the YOLOv3 model to locate, classify, and spray weeds with 96% accuracy under field conditions. The literature broadly agrees that Recall, F-1 Score, and AUC, together with Precision, constitute the most commonly adopted yardsticks for gauging model accuracy. In our study, we found these performance indicators to be tightly intertwined with Precision. A positive linear relationship was detected between Precision and Recall ($R^2 = 0.98$), Precision and F-1 Score ($R^2 = 0.99$), and Precision and AUC ($R^2 = 0.95$). Consequently, leaning on merely one or two of these metrics could conceivably yield a faithful appraisal of the model. Our results also highlighted that, by fielding the YOLOv5 small neural network, all three weed species could be recognized at every growth stage, with a precision of 0.86, except at the fruit stage. Turning to the medium, large, and xlarge architectures diminished weed detection precision by 8%, 7%, and 8%, respectively. This suggests that passable weed detection can be achieved with YOLOv5 flavours other than small, albeit at an accuracy penalty of around 8%. Stacked against medium, large, and xlarge, opting for nano triggered a 21% erosion in weed species prediction precision, a strong sign that nano should be bypassed when the goal is detecting the three chosen weed species in wheat fields.

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

When examining the output of the trained deep learning architectures for weed detection, it becomes clear that they deliver high accuracy, ranging from 89% to 97%, for the first four phenological stages—those of greatest consequence for weed management and suppression across the three weed species. Spotting a weed during the closing chapter of its life cycle, the fruiting stage, furnishes valuable feedback on the efficacy of applied agricultural control tactics and opens the door to predicting probable locations for the next growth cycle. Even though the fifth phenological stage of forking larkspur in the field was captured with a 45% success rate, the terminal stages of the remaining two weed species (charlock mustard and creeping thistle) were correctly forecast with rates spanning 81% to 88%. The competence demonstrated by these trained deep learning architectures can serve as a formidable tool for advanced agricultural practice.

This investigation carried multiple practical ramifications. For instance, its outcomes showed that all three target weed species could be flagged with high accuracy at every stage, except fruiting. This opens the door to more precisely directed integrated weed management and control measures. Growers should emphasize the precision placement of herbicides to reduce off-target effects and prevent environmental pollution. With knowledge of resident weed communities, practitioners can distinguish and sort various weeds, classifying them by phenological stage. Training ANNs with imagery captured by multispectral and hyperspectral sensors might yield better results when separating objects that are difficult to tease apart using RGB spectral hues alone. As the roster of weed species identifiable through artificial intelligence lengthens, the odds of reliably discriminating each weed from the next likewise improve. Putting computer vision to work for weed detection and control promises to lift crop yields, delivering economic dividends to farming enterprises.

Conclusion

We concluded that YOLOv5s (small) was the most precise neural network for weed detection across species. Our results indicated that YOLOv5n (nano) should not be enlisted for weed detection, whereas YOLOv5 medium, large, and xlarge returned satisfactory results, though with an 8% accuracy deficit relative to YOLOv5s. Weed detection using the small neural network yielded accuracy rates that varied by weed species. For CT, detection precision varied from 0.87 (fruit stage) to 0.96 (rosette and vegetative stages). For forking larkspur, the optimal neural network performance was achieved with YOLOv5 small, and precision ranged from 0.45 (fruit stage) to 0.93 (vegetative stage). In Charlock mustard, the most precise detection occurred at the seedling, rosette, and vegetative stages (precision of 0.96 each), while the least precise occurred at the fruit stage (precision of 0.81). We recommend deploying the YOLOv5 small neural network to detect all three weed species across all growth stages, achieving a precision of 0.86, except during the fruit stage. Upcoming studies should (1) prioritize benchmarking the YOLOv5 small neural network against the freshest iterations, such as YOLOv7, and also (2) assess this neural network for flagging weed species in wheat fields to support enhanced weed control and shape future guidance. In addition, this study leaned on 2D pixel-based features pulled from drone-captured RGB imagery for object identification and detection. Future inquiries could juxtapose these outcomes with 3D reconstructions generated by software packages such as Pix4D. Furthermore, emerging technologies such as LIDAR, multispectral, and hyperspectral sensors could provide richer, more discriminating information for distinguishing objects. These technologies warrant exploration in follow-up investigations to improve the accuracy and efficiency of object detection and identification.

The single most substantial contribution of our work was the assembly of a distinctive dataset that stands apart from the existing body of published literature. This dataset was tested with just one deep learning model, yet its promise extends far beyond that. It can be tapped for a variety of studies that employ different deep learning models and libraries—both those available today and those that will emerge in the future. In doing so, a reference point can be established for measuring the performance of diverse deep learning architectures on a dataset that has not previously been used in deep learning object detection research. Future research should pivot toward alternative deep learning methods and the re-examination of the assembled dataset by other investigators, enabling a comparison of their findings with ours. This kind of comparative assessment will yield rich insights into the flexibility and potency of different deep learning strategies for

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

weed detection when applied to our dataset. Such a collaborative, iterative process holds substantial promise for advancing weed detection and sharpening the accuracy and real-world applicability of future detection models.

Acknowledgements

The article emerged and was revised by the PhD thesis “Detection of some weeds in wheat production of Tokat region by deep learning,” defended at Tokat Gaziosmanpasa University, Tokat, 2022. We want to thank Dr. Khaled R. Ahmed for his generous help and for sharing his valuable experiences.

Conflict of interest

None

Financial support

This study was supported by the TUBITAK (The Scientific and Technological Research Council of Türkiye) (3501-120O888).

Ethics statement

None

Received: 09 Mar 2023 Revised: 14 May 2023 Accepted: 30 May 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. Cramer H. Pflanzenschutz und Welternte. Bayer Leverkusen: Pflanzenschutz Nachrichten; 1967.
2. Özer Z, Kadıoğlu İ, Önen H, Tursun N. Herboloji (Yabancı ot bilimi). Tokat: Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Yayınları; 2003. p. 579.
3. Acıbuca V. Economic analysis of durum wheat production in Mardin. Adana: Cukurova University; 2010.
4. Atak M. Farklı triticale hatlarının morfolojik ve DNA markörleriyle genetik karakterizasyonu. Ankara: Ankara Üniversitesi Fen Bilimleri Enstitüsü; 2004.
5. Aydın N, Bayramoğlu H, Mut Z, Özcan H. Ekmeklik buğday (Triticum aestivum L.) çeşit ve hatlarının karadeniz koşullarında verim ve kalite özelliklerinin belirlenmesi. AÜZF Tarım Bilimleri Dergisi. 2005;11(3):257-62.

6. Combellak J. Weeds in cropping their cost to the Australian economy. Plant Prot Q. 1987;2(2).

7. Oerke E, Dehne H, Schönbeck F, Weber A. Crop production and crop protection: Estimated losses in major food and cash crops. Elsevier; 1994.

8. Jabran K, Mahajan G, Sardana V, Chauhan B. Allelopathy for weed control in agricultural systems. Crop Prot. 2015;72:57-65.

9. Gharde Y, Singh PK, Dubey RP, Gupta PK. Assessment of yield and economic losses in agriculture due to weeds in India. Crop Prot. 2018;107:12-8.

10. Oerke EC. Crop losses to pests. J Agric Sci. 2006;144(1):31-43.

11. Chauhan BS. Grand challenges in weed management. Front Agron. 2020;1(3).

12. Sin B, Kadioglu İ. Trp-574-Leu mutation in wild mustard (Sinapis arvensis L.) as a result of ALS inhibiting herbicide applications. PeerJ. 2021;9:e11385.

13. Mennan H. Determination of weed species seen in wheat cultivation areas of Samsun province and investigation of germination and development biology of important species. Adana: Cukurova University; 1993.

14. Mennan H, Uygur F. Detection of weeds in wheat cultivation areas of Samsun province. J Agric Fac Ondokuz Mayıs Univ. 1994;9:25-35.

15. Boz Ö. Determination of weed species, frequency and density in wheat growing areas in Aydın. Turk J Weed Sci. 2000;3(2):1-11.

16. Tursun N. Determination of weeds in wheat fields in Kahramanmaraş, Turkey. Turk J Weed Sci. 2002;5(1):1-11.

17. Üstüner T, Güncan A. A research on weed species which are problem, importance, biology of germination and control possibilities of them in potato fields in Niğde province. Konya: Department of Plant Protection; 2002.

18. Kitiş Y, Boz Ö. Determination of the weed species, their observation frequencies and densities, in wheat fields in Isparta province. Turk J Weed Sci. 2003;6(1):24-5.

19. Mennan H, Işık D. Germination ability and densities of weed seeds contamination in wheat seeds. Turk J Weed Sci. 2003;6(1):8-15.

20. Gökalp Ö, Üremiş İ. Mardin buğday ekim alanlarında bulunan yabancı ot türlerinin, yaygınlıklarının ve yoğunluklarının belirlenmesi. Mustafa Kemal Üniversitesi Ziraat Fakültesi Dergisi. 2015;20(1):13-22.

21. Gürsu Z. Determination of significant weed species in wheat growing areas of Kırklareli province, their intense and frequency of coincidance. Tekirdağ: Namık Kemal University; 2015.

22. Şin B, Kadıoğlu İ, Kamışlı B. Determination of weed seeds in wheat grain in Tokat Province. Turk J Weed Sci. 2016;19(2):28-37.

23. Pala F, Mennan H, Çığ F, Dilmen H. Determination of weed seeds mixed with wheat product in Diyarbakır. Turk J Agric Res. 2018;5(3):183-90.

24. Boz Ö, Uygur F. Çukurova Bölgesi buğday ekim alanlarındaki yabani hardal (Sinapis arvensis L.) ve yabani fiğin (Vicia spp.) zarar seviyelerinin saptanması ve ekonomik zarar eşiğinin hesaplanması. Türkiye II. Herboloji Kong. 1997:1-4.

25. Özer Z, Önen H, Tursun N, Uygur F. Türkiye'nin bazı önemli yabancı otları (Tanımları ve kimyasal savaşımları). Gaziosmanpaşa Üniversitesi Ziraat Fakültesi Yayınları Kitap Serisi. 1999;38(16).

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

26. El-Faki MS, Zhang N, Peterson DE. Weed detection using color machine vision. Trans ASAE. 2000;43(6):1969-78.

27. Gerhards R, Christensen S. Real-time weed detection, decision making and patch spraying in maize, sugarbeet, winter wheat and winter barley. Weed Res. 2003;43(6):385-92.

28. Zhu W, Zhu X. The application of support vector machine in weed classification. IEEE International Conference on Intelligent Computing and Intelligent Systems. 2009.

29. Shapira U, Herrmann I, Karnieli A, Bonfil DJ. Field spectroscopy for weed detection in wheat and chickpea fields. Int J Remote Sens. 2013;34(17):6094-108.

30. Bakhshipour A, Jafari A, Nassiri SM, Zare D. Weed segmentation using texture features extracted from wavelet sub-images. Biosyst Eng. 2017;157:1-12.

31. Kamilaris A, Prenafeta-Boldu F. Deep learning in agriculture: A survey. Comput Electron Agric. 2018;147:70-90.

32. Tian H, Wang T, Liu Y, Qiao X, Li Y. Computer vision technology in agricultural automation – a review. Inf Process Agric. 2020;7(1):1-19.

33. Guzel M, Sin B, Turan B, Kadioglu I. Real-time detection of wild mustard (Sinapis arvensis L.) with deep learning (YOLO-v3). Fresenius Environ Bull. 2021;30(11/A):12197-203.

34. Bock CH, Poole GH, Parker PE, Gottwald TR. Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging. Crit Rev Plant Sci. 2015;29:59-107.

35. Bagheri A, Heydari Z. Classification of weed seeds using image processing and pattern recognition methods. Iran J Weed Sci. 2021;17(2):65-77.

36. Rumpf T, Römer C, Weis M, Sökefeld M, Gerhards R, Plümer L. Sequential support vector machine classification for small-grain weed species discrimination with special regard to Cirsium arvense and Galium aparine. Comput Electron Agric. 2012;80:89-96.

37. Pantazi X-E, Moshou D, Bravo C. Active learning system for weed species recognition based on hyperspectral sensing. Biosyst Eng. 2016;146:193-202.

38. de Castro AI, López-Granados F, Jurado-Expósito M. Broad-scale cruciferous weed patch classification in winter wheat using QuickBird imagery for in-season site-specific control. Precision Agric. 2013;14:392-413.

39. Rasmussen J, Nielsen J, Streibig JC, Jensen JE, Pedersen KS, Olsen SI. Pre-harvest weed mapping of Cirsium arvense (L.) Scop in wheat and barley with off-the-shelf UAVs. Precision Agric. 2019;20(5):983-99.

40. Sánchez-Sastre LF, Casterad MA, Guillén M, Ruiz-Potosme NM, Veiga NMAD, Navas-Gracia LM, et al. UAV detection of Sinapis arvensis infestation in alfalfa plots using simple vegetation indices from conventional digital cameras. AgriEngineering. 2020;2(2):206-12.

41. Özgür A, Nar F. Derin öğrenme yöntemleri kullanarak ekin ile yabancı otların birbirinden ayırt edilmesi. International Conference on Computer Technologies and Applications in Food and Agriculture. Konya. 2019:76-89.

42. Ghersa CM, Holt JS. Using phenology prediction in weed management: A review. Weed Res. 1995;35(6):461-70.

43. Midtiby HS, Åstrand B, Jørgensen O, Jørgensen RN. Upper limit for context–based crop classification in robotic weeding applications. Biosyst Eng. 2016;146:183-92.

44. Deng X, Qi L, Ma X, Jiang Y, Chen X, Liu H, et al. Recognition of weeds at seedling stage in paddy fields using multi-feature fusion and deep belief networks. Trans Chin Soc Agric Eng. 2018;34(14):165-72.

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

45. Sharma P. Crops and weeds classification using Convolutional Neural Networks via optimization of transfer learning parameters. Int J Eng Adv Technol. 2019:2249-8958.

46. Fawakherji M, Youssef A, Bloisi D, Pretto A, Nardi D. Crop and weeds classification for precision agriculture using context-independent pixel-wise segmentation. 2019 Third IEEE International Conference on Robotic Computing. IEEE. 2019.

47. Koirala A, Walsh K, Wang Z, McCarthy C. Deep learning method overview and review of use for fruit detection and yield estimation. Comput Electron Agric. 2019;162:219-34.

48. Elstone L, How KY, Brodie S, Ghazali MZ, Heath WP, Grieve B. High speed crop and weed identification in lettuce fields for precision weeding. Sensors. 2020;20(2):455.

49. Leminen Madsen S, Mathiassen SK, Dyrmann M, Laursen MS, Paz LC, Jørgensen RN. Open plant phenotype database of common weeds in Denmark. Remote Sens. 2020;12(8):1246.

50. Sudars K, Jasko J, Namatevs I, Ozola L, Badaukis N. Dataset of annotated food crops and weed images for robotic computer vision control. Data Brief. 2020;31:105833.

51. Champ J, Mora-Fallas A, Goëau H, Mata-Montero E, Bonnet P, Joly A. Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. Appl Plant Sci. 2020;8(7):e11373.

52. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436-44.

53. Caliskan A, Badem H, Basturk A, Yuksel ME. Classification and diagnosis of cardiac arrhythmia disease by deep learning. International Conference on Artificial Intelligence and Data Processing (IDAP16). 2016.

54. Badem H, Caliskan A, Basturk A, Yuksel ME. Classification and diagnosis of the Parkinson disease by stacked autoencoder. 10th International Conference on Electrical and Electronics Engineering ELECO. 2016.

55. Jocher G, Stoken A, Borovec J, Changyu L, Hogan A, Diaconu L, et al. Ultralytics/yolov5: v3.0. Zenodo. 2020.

56. Marom ND, Rokach L, Shmilovici A. Using the confusion matrix for improving ensemble classifiers. IEEE 26th Convention of Electrical and Electronics Engineers in Israel. 2010:555-9.

57. Metz CE. Basic principles of ROC analysis. Semin Nucl Med. 1978:283-98.

58. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology. 1982;143(1):29-36.

59. Altman DG, Bland JM. Statistics notes: diagnostic tests 2: predictive values. BMJ. 1994;309(6947).

60. Fawcett T. An introduction to ROC analysis. Pattern Recognit Lett. 2006;27(8):861-74.

61. Rani KA, Supriya P, Sarath TV. Computer vision-based segregation of carrot and curry leaf plants with weed identification in carrot field. International Conference on Computing Methodologies and Communication (ICCMC). 2017:185-8.

62. Ishak AJ, Hussain A, Mustafa MM. Weed image classification using Gabor wavelet and gradient field distribution. Comput Electron Agric. 2009;66(1):53-61.

63. He D, Qiao Y, Li P, Gao Z, Li H, Tang J. Weed recognition based on SVM-DS multi-feature fusion. Nongye Jixie Xuebao (Trans Chin Soc Agric Mach). 2013;44(2):182-7.

64. Herrmann I, Shapira U, Kinast S, Karnieli A, Bonfil DJ. Ground-level hyperspectral imagery for detecting weeds in wheat fields. Precision Agric. 2013;14(6):637-59.

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

65. Haug S, Ostermann J. A Crop/Weed Field Image Dataset For the Evaluation of Computer Vision-Based Precision Agriculture Tasks. Computer Vision–ECCV 2014 Workshops. Springer International Publishing. 2014:105-16.

66. Chaki J, Parekh R, Bhattacharya S. Plant leaf recognition using texture and shape features with neural classifiers. Pattern Recognit Lett. 2015;58:61-8.

67. Murad NY, Mahmood T, Forkan ARM, Morshed A, Jayaraman PP, Siddiqui MS. Weed detection using deep learning: A systematic literature review. Sensors. 2023;23(7):3670.

68. Giselsson TM, Jørgensen RN, Jensen PK, Dyrmann M, Midtiby HS. A public image database for benchmark of plant seedling classification algorithms. arXiv preprint arXiv:1711.05458. 2017.

69. Dos Santos Ferreira AFD, da Silva GG, Pistori H, Folhes MT. Weed detection in soybean crops using ConvNets. Comput Electron Agric. 2017;143:314-24.

70. Bakhshipour A, Jafari A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. Comput Electron Agric. 2018;145:153-60.

71. Chavan TR, Nandedkar AV. AgroAVNET for crops and weeds classification: A step forward in automatic farming. Comput Electron Agric. 2018;154:361-72.

72. Sabzi S, Abbaspour-Gilandeh Y, Arribas J. An automatic visible-range video weed detection, segmentation and classification. Heliyon. 2020;6(5).

73. Junior LCM, Ulson JAC. Real time weed detection using computer vision and deep learning. 14th IEEE International Conference on Industry Applications (INDUSCON). 2020.

74. Ruigrok T, van Henten E, Booij J, Van Boheemen K, Kootstra G. Application-specific evaluation of a weed-detection algorithm for plant-specific spraying. Sensors. 2020;20(24):7262.

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