COMPARATIVE YOLOV8–YOLOV9 PERFORMANCE IN WEED RECOGNITION FOR COTTON FIELDS

Uteuliev N.U. [1[0009-0003-0698-5635]], **Dauletnazarov J.I.** [2[0009-0005-7108-2060]]

[1]Head of the Department of Software Engineering at Nukus State Technical University
[2]Tashkent University of Information Technologies named after Muhammad al-Khwarizmi
E-mail: dauletnazarovjaksilik@gmail.com

Annotation. This research utilizes a proprietary dataset consisting of images captured from cotton fields in the Republic of Karakalpakstan. A new annotated image dataset from cotton fields in the Nukus district was developed to compare the performance of YOLOv8 and YOLOv9 models in multi-class weed detection. YOLOv9c has outperformed the competing model YOLOv8m with respect to overall accuracy (mAP@0.5 = 0.865). YOLOv9c is therefore more appropriate for ongoing precise weed management during real-time applications. YOLOv8m, however, may still be deployed on energy-efficient hardware.
Keywords: YOLOv8, YOLOv9, weed detection, cotton production, deep learning, computer vision, mAP, precision agriculture, image dataset.

Annotatsiya. Ushbu tadqiqot Qoraqalpog'iston Respublikasidagi paxta dalalaridan olingan tasvirlardan iborat mulkiy ma'lumotlar to'plamidan foydalanadi. Nukus tumanidagi paxta dalalaridan olingan yangi izohli tasvir ma'lumotlar to'plami YOLOv8 va YOLOv9 modellarining ko'p sinfli begona o'tlarni aniqlashdagi samaradorligini taqqoslash uchun ishlab chiqilgan. YOLOv9c umumiy aniqlik bo'yicha raqobatchi YOLOv8m modelidan ustun keldi (mAP@0.5 = 0.865). Shuning uchun YOLOv9c real vaqt rejimida qo'llaniladigan dasturlarda begona o'tlarni aniq boshqarish uchun ko'proq mos keladi. Biroq, YOLOv8m hali ham energiya tejaydigan uskunalarda qo'llanilishi mumkin.
Kalit so'zlar: YOLOv8, YOLOv9, begona o'tlarni aniqlash, paxta ishlab chiqarish, chuqur o'rganish, kompyuter ko'rish, mAP, aniq qishloq xo'jaligi, tasvir ma'lumotlar to'plami.

Аннотация. В данном исследовании используется собственный набор данных, состоящий из изображений хлопковых полей в Республике Каракалпакстан. Для сравнения производительности моделей YOLOv8 и YOLOv9 в многоклассовом обнаружении сорняков был разработан новый аннотированный набор данных изображений хлопковых полей Нукусского района. Модель YOLOv9c превзошла конкурирующую модель YOLOv8m по общей точности (mAP@0.5 = 0.865). Таким образом, YOLOv9c больше подходит для точного управления сорняками в режиме реального времени. Однако YOLOv8m может быть развернута на энергоэффективном оборудовании.
Ключевые слова: YOLOv8, YOLOv9, обнаружение сорняков, производство хлопка, глубокое обучение, компьютерное зрение, mAP, точное земледелие, набор данных изображений.

Introduction

Weeds take up space, nutrients, water, and light that would otherwise help cotton grow well. As a result, cotton growers will frequently lose money on speed of harvest and produce less cotton than expected [1].
Recent progress in deep learning and computer vision gives rise to new opportunities for field-based automation of weed identification and categorization. Across the broad landscape of computer vision methodologies, detection-oriented neural architectures particularly the lineage derived from the YOLO (You Only Look Once) paradigm have emerged as a prominent research focus[2][3].
Prior studies have demonstrated the feasibility of applying YOLO-based models for multi-class weed detection in cotton fields. For example, the dataset presented in CottonWeedDet12 contains 5,648 images collected under natural field conditions and annotated with 9,370 bounding boxes covering 12 common weed species in cotton production systems [1]. Similarly, recent works such as YOLO-WDNet and Cotton

Weed-YOLO build upon YOLO architectures to improve detection accuracy while reducing computational load - a critical requirement for real-world agricultural applications [4, 5].

Many existing datasets and studies were conducted in agro-climatic zones different from Central Asia, and they may not fully reflect the weed species composition, soil background, illumination conditions, and crop management practices typical for regions such as Uzbekistan. As well, weed species frequently display significant morphological inequality during growth phases, from seedling to mid-grown, leading to proper identification in the field more challenging. One recent study found that though models were trained on a variety of images from several years, model performance diminished significantly when attempting to classify multiple growth stages of Amaranthus palmeri (i.e., Palmer Amaranth) within cotton. This significant decrease indicates the considerable difficulty in accurately generalizing across multiple phenological changes [6].

As a result, it proves crucial to develop and evaluate datasets and detection models that are unique to local conditions. In this work, we collected a region-specific dataset from the cotton-growing fields of the Nukus district (Republic of Karakalpakstan). Our sampling took place throughout two farming seasons, following plants as they grew so we could observe realistic changes in both their form and surroundings. Given the tradeoffs between accuracy, computational complexity, and real-time inference speed, we selected two mid-sized YOLO variants - YOLOv8m and YOLOv9c - to conduct a comparative analysis under identical training conditions.

The main purpose of this study is to collect and label images of cotton, the most common weeds, and other plants found in cotton fields in Central Asia particularly in Uzbekistan, test how well YOLOv8 and YOLOv9 can recognize different types of weeds in real field conditions, see if these models can work fast enough for real time use in the field, such as on farming robots or smart equipment; and find the current challenges and suggest how the dataset and models can be improved in the future.

Methodology

RGB images of plants were colleted from cotton fields in Nukus district, Rebuplic of Karakalpakstan in two seasons. The first collection campaign took place in April-May 2024, approximately 15-20 days after cotton sowing, resulting in 225 RGB images. Then, to increase the number of images in the dataset 856 more pictures of cotton and weeds were collected in 2025. This time pictures were taken in different growth stages of cotton (June-August) to fully capture morphological and physiological features of the plants.This two-season sampling strategy allowed representation of early vegetative to mid-growth phases of both cotton and weed species. Images were captured using a Xiaomi 11 Lite 5G NE smartphone (6944×9280 px resolution). To simulate typical viewpoints for mobile agricultural robots, the camera was positioned at heights of 50 cm and 100 cm with 80-90° viewing angles relative to the soil surface. All images were recorded in JPG format. After taking a close look at image quality, we filtered out any that lacked clarity, and ultimately arrived at a refined dataset of 1,081 strong, usable images.

Target classes and annotation procedure. Our dataset focuses on eight biologically and agriculturally meaningful classes encompassing cotton (Gossypium spp.) and seven major weed species widely distributed across the Republic of Karakalpakstan's cotton-producing fields, which makes them critical priorities for automated identification. Annotations were performed manually on the Roboflow platform using rectangular bounding boxes. The result of this careful quality assurance process is a dataset comprising 4,411 rigorously validated objects.

The specific taxonomy and quantitative distribution of instances per class are detailed in Table 1.

Table 1.

Distribution of labeled instances across target classes.

Common name	Scientific name	Count
Cotton	Gossypium spp.	2,177
Large Crabgrass	Digitaria sanguinalis	1,385
Common Reed	Phragmites australis	228
Field Bindweed	Convolvulus arvensis	162
Cocklebur	Xanthium strumarium	151
Lambsquarters	Chenopodium album	128
Velvetleaf	Abutilon theophrasti	70
Other weeds	Weed spp. (Mixed)	110
Total		4,411

In real fields, weed emergence is rarely orderly some patches are dense with weeds while others have almost none.

Preprocessing and dataset partitioning. We reformatted each image to 640x640 pixels while maintaining its original geometry with aspect-ratio-aware padding in order to strike a balance between processing requirements and visual accuracy. Whereas an beginning arbitrary part was connected, stratification was afterward presented to adjust regular variety and phenological differences over subsets. The final dataset partition was: 1) Training set: 744 images. 2) Validation set: 167 images. 3) Test set: 170 images.

Data augmentation

We adopted a two-step augmentation strategy to prepare the model for the unpredictable nature of field imagery, helping it recognize plants accurately even when circumstances differ from the training data. Crucially, all geometric augmentations (rotation, translation, scaling, mosaic, copy-paste) were applied consistently to both the image pixel data and the corresponding bounding box coordinates to maintain label accuracy.

Augmentation before training process (Pre-training). Applied via the Roboflow platform to expand the base dataset before training commences. Pre-export image preparation modifications are summarized in Table 2.

Table 2.

Pre-export image preparation modifications.

Modification	Description
Rotations at 90° increments	Images were rotated at 90 increments (clockwise, counter-clockwise, upside down).
Random rotation	Random rotation between –13 and +13.
Hue adjustment	Hue adjusted in a range from –23 to +23.
Saturation	Saturation adjusted between –30% and +30%.

adjustment	
Brightness enhancement	Brightness enhanced between 0–15%.

Augmentation during the training process included dynamically by the Ultralytics YOLO engine during the training pipeline:
Mosaic (enabled, disabled last 10 epochs)
Horizontal flip (p = 0.5)
HSV augmentation (h = 0.015, s = 0.7, v = 0.4)
Translation = 0.1
Scaling = 0.5
Random erasing = 0.4
Copy-paste augmentation
Randaugment
Both augmentation strategies were applied uniformly across all tested YOLO versions (YOLOv8: n, s, m, l, x; YOLOv9: c, e) to ensure a fair comparative analysis.
All models were trained using the Ultralytics YOLO framework on a high-performance workstation tailored for deep learning tasks. To handle the demanding nature of deep learning tasks, we used a powerful workstation built for AI research. In our lab to train our models we used computer with this specification an Intel Core i9-12900K CPU and 128 GB of DDR5 RAM with a Gigabyte GeForce RTX 4090 GPU (24 GB VRAM), giving us the processing strength needed for our experiments. High-speed storage was delivered by two 2 TB Samsung 980 PRO NVMe SSDs, which kept data loading quick and helped minimize waiting between training runs.
Each YOLO model was trained under identical hyperparameters to ensure fair evaluation, the information about hyperparameters is detailed in Table 3.

Table 3.

Hyperparameters for training YOLOv8 and YOLOv9 algorithms

Hyperparameter	Value
Epochs	150
Batch size	16 (YOLOv9c used 32)
Image size	640 × 640
Optimizer	AdamW
Learning rate	0.001
Pretrained weights	Yes
Mixed precision	Enabled
Early stopping patience	50 epochs

Evaluation metrics. The performance of the proposed object detection models was evaluated using standard metrics, namely precision, recall, and mean Average Precision (mAP) [1, 2]. In straightforward terms, accuracy answers the address, of everything the demonstrate stamped as positive, how numerous were rectify (1).

Review answers, of all the positives that exist, how numerous did the demonstrate discover (2). Both depend on genuine positives as the numerator but vary in what they compare against. The mean Average Precision (mAP) is calculated as the mean of the average precision (AP) values across all object classes (3).

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

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

$$mAP = (\frac{1}{N}) \times \sum(AP_i)_{i=1}{}^{N} \quad (3)$$

Results and Discussion

Table 4 summarizes the global detection performance of the YOLOv8 and YOLOv9 model families for multi-species weed detection. Looking across the YOLOv8 variants, precision was consistently high (0.804–0.849), recall followed closely (0.709–0.797), and mAP@0.5 results also remained strong at 0.790–0.826. YOLOv8m achieved the best overall balance, with precision of 0.837, recall of 0.797, and the highest mAP@0.5 (0.826) as well as the highest mAP@[0.5:0.95] (0.634) among the YOLOv8 variants. Despite its enhanced expressive capability leading to a slightly superior mAP@[0.5:0.95] value of 0.612, YOLOv8x did not transcend YOLOv8m in overall accuracy. Strong performance in actual agricultural settings is largely dependent on how well the models architecture handles data since this is crucial for enabling the system to adapt to various field conditions without sacrificing accuracy.

The YOLOv9 models outperformed YOLOv8 in almost all global metrics. Relative to YOLOv8m, YOLOv9c exhibited notable performance enhancements, attaining 0.880 precision, 0.832 recall, and 0.865 mAP@0.5, which constitute respective increments of 3.8, 3.5, and 3.9 percentage points. In contrast, YOLOv9e achieved a higher mAP@[0.5:0.95] of 0.663, reflecting improved multi-threshold localization behavior. These improvements are consistent with recent studies where enhanced YOLOv8 architectures for cotton weed detection also surpassed baseline YOLO variants in terms of mAP and F1 score while remaining suitable for field deployment [10, 11, 12].

YOLOv8m and YOLOv9c were deemed the most suitable for deeper evaluation because they deliver a strong speed-accuracy balance relative to other configurations. Looking ahead, we plan to extend this research by testing the latest YOLO versions and validate their operational effectiveness through on-site field trials.

Table 4.

Evaluation metrics of YOLOv8 and YOLOv9 architectures for weed detection

YOLO models		Precision	Recall	mAP@0.5	mAP@[0.5:0.95]
YOLOv8	YOLOv8n	0.826	0.709	0.790	0.583
	YOLOv8s	0.849	0.719	0.809	0.579
	YOLOv8m	0.837	0.797	0.826	0.634
	YOLOv8l	0.804	0.721	0.817	0.595
	YOLOv8x	0.846	0.749	0.815	0.612
YOLOv9	YOLOv9c	0.880	0.832	0.865	0.640
	YOLOv9e	0.825	0.785	0.849	0.663

References

[1]. F. Dang, D. Chen, Y. Lu, Z. Li. YOLOWeeds: A novel benchmark of YOLO object detectors for multi-class weed detection in cotton production systems. Journal: "Computers and Electronics in Agriculture", (205) 107655, 2023.

[2]. K. S. Kamalesh, R. Kumaraperumal, P. Pazhanivelan, R. Jagadeeswaran, P. C. Prabu. YOLO deep learning algorithm for object detection in agriculture: A review. Journal: "Journal of Agricultural Engineering". 2024. DOI: https://doi.org/10.4081/jae.2024.1641

[3]. A. Das, Y. Yang, V.H. Subburaj. YOLOv7 for weed detection in cotton fields using UAV imagery. Journal: "AgriEngineering". https://doi.org/10.3390/agriengineering7100313

[4]. J. Hu, H. Gong, S. Li, Y. Mu, Y. Guo, Y. Sun, T. Hu, Y. Bao. Cotton Weed-YOLO: A lightweight and highly accurate cotton weed identification model for precision agriculture. Journal: "Agronomy".2025. DOI: https://doi.org/10.3390/agronomy14122911

[5]. X. Fan, T. Sun, X. Chai, J. Zhou. YOLO-WDNet: A lightweight and accurate model for weeds detection in cotton field. Journal: "Computers and Electronics in Agriculture". 2024. DOI: https://doi.org/10.1016/j.compag.2024.109317

[6]. G. R. Coleman, M. Kutugata, M. J. Walsh, M. Bagavathiannan. Multi-growth stage plant recognition: A case study of Palmer amaranth (Amaranthus palmeri) in cotton (Gossypium hirsutum). Journal: "arXiv preprint". 2023. DOI: https://doi.org/10.48550/arXiv.2307.15816

[7]. A. K. Kanade, R. K. Patel, M. J. Shah. Weed detection in cotton farming by YOLOv5 and YOLOv8 object detectors. Journal: "Biosystems Engineering". 2025. DOI: https://doi.org/10.1016/j.biosystemseng.2025.06.012

[8]. M. J. Karim, M. Nahiduzzaman, M. Ahsan, J. Haider. Development of an early detection and automatic targeting system for cotton weeds using an improved lightweight YOLOv8 architecture on an edge device. Journal: "Knowledge-Based Systems". 2024. DOI: https://doi.org/10.1016/j.knosys.2024.112204

[9]. C. Ma, G. Chi, X. Ju, J. Zhang, C. Yan. YOLO-CWD: A novel model for crop and weed detection based on improved YOLOv8. Journal: "Crop Protection". 2025. DOI: https://doi.org/10.1016/j.cropro.2025.107169

[10]. D. Ren, W. Yang, Z. Lu, D. Chen, H. Shi. Improved weed detection in cotton fields using enhanced YOLOv8s with modified feature extraction modules. Journal: "Symmetry". 2024. DOI: https://doi.org/10.3390/sym16040450

[11]. J. Wang, Z. Qi, Y. Wang, Y. Liu. A lightweight weed detection model for cotton fields based on an improved YOLOv8n. Journal: "Scientific Reports". 2025. DOI: https://doi.org/10.1038/s41598-024-84748-8

[12]. L. Zheng, J. Yi, P. He, J. Tie, Y. Zhang, W. Wu, L. Long. Improvement of the YOLOv8 model in the optimization of the weed recognition algorithm in cotton field. Journal: "Plants". 2024. DOI: https://doi.org/10.3390/plants13131843

[13]. L. Zheng, C. Zhu, L. Liu, Y. Yang, J. Wang, W. Xia, K. Xu, J. Tie. Star-YOLO: A lightweight and efficient model for weed detection in cotton fields using advanced YOLOv8 improvements. Journal: "Computers and Electronics in Agriculture". 2025. DOI: https://doi.org/10.1016/j.compag.2025.110306