Towards the Development of a Deep Learning Based Maize-Weed Detection System using YOLO Architectures

Taliha Abiodun Folorunso
Department of Mechatronics Engineering
Federal University of Technology, Minna
Minna, Nigeria
funso.taliha@futminna.edu.ng

Jibril Abdullahi Bala
Department of Mechatronics Engineering
Federal University of Technology, Minna
Minna, Nigeria
Jibril.bala@futminna.edu.ng

Trusting Inekwe
Data Science and Analytics Department, & Artificial Intelligence Exploration Center,
SUNY Polytechnic Institute,
Utica, New York, USA
inekwet@sunypoly.edu

Adegboyega Adebayo
African Centre of Excellence on Technology Enhanced Learning,
National Open University of Nigeria
Abuja, Nigeria
aadebayo@noun.edu.ng

Anjana Ran Yadav
Artificial Intelligence Exploration Center,
SUNY Polytechnic Institute, Utica, New York, USA
anjanayadav55555@gmail.com

William Thistleton
Artificial Intelligence Exploration Center,
SUNY Polytechnic Institute, Utica, New York, USA
thistlet@sunypoly.edu

***Abstract*—Weeds are a primary constraint to agricultural productivity, potentially reducing crop yields by 30% and making crop production challenging. Conventional and chemical management methods are often labour-intensive or environmentally damaging; thus, there is a need for a better approach that leverages advances in deep learning techniques. This study investigates the performance of nine YOLO variants (v8–v12) in nano and small sizes for real-time maize-weed detection in precision agriculture. The evaluation shows a significant trade-off between computational efficiency and detection robustness. While the newer YOLOv12s architecture is more optimised, with 21.2 GFLOPs, the older YOLOv8s maintains a performance lead, achieving a peak mAP@50 of 0.870 and a recall of 0.861. The results indicate that YOLOv8n achieves the fastest inference time (4.4 ms), making it well-suited for high-speed monitoring that requires maximum weed eradication, given its high recall. Finally, YOLOv12n offers the best balance for edge-based devices, given its size, thereby demonstrating the capability of the lightweight Yolo model to be deployed for weed identification and detection in maize farmlands.**

Keywords—Weed detection, Crop Detection, Machine Learning, Deep Learning, YOLO

I. INTRODUCTION

Agriculture is essential for global food security, feeding the world's ever-increasing population of billions of people [1], [2]. It provides essential crops, sustains livelihoods, and underpins economic development [3]. However, weeds represent the most significant biotic constraint to agricultural productivity worldwide, causing more yield loss and adding greater production costs than insect pests, pathogens, or nematodes [4]. Weeds compete aggressively with crops for water, sunlight, nutrients, and space, potentially reducing crop yields by up to 30% globally or causing complete crop failure if left uncontrolled [5], [6]. They also harbour pests and diseases, contaminate harvests, and complicate agricultural operations.

To control the influence and impact of weeds on farmland and agricultural production, several approaches have been implemented, including Conventional, Chemical, and Technological approaches. Conventional methods involve manually removing unwanted weeds from farmland using agricultural tools[7]. However, this is not feasible in very large-scale farming settings, as it is laborious, time-consuming, and resource-intensive. Chemical approaches involve the use of herbicides and pesticides for the management of weeds [7], [8], [9], [10]. However, challenges arise in the selective application of these chemicals, as they can affect parent crops in multi-crop settings. In the use of technology, modern farm machinery has been developed to support farm management practices in weed removal [11], [12].

Furthermore, with advances in technology and the emergence of Precision Agriculture, more robust approaches to identifying and managing weeds in farmlands have been developed [12]. These approaches range from embedded systems to Artificial Intelligence (AI). AI approaches leverage the feature-extraction and pattern-recognition capabilities of Machine Learning (ML) and Deep learning (DL) algorithms to identify and detect weeds and their parent crops. Traditional ML, DL, and computer vision approaches have been inadequate for identifying these weeds due to their limitations in efficiency, performance, and resource requirements [2],[7]. However, with advancements in algorithm development, Convolutional Neural Networks (CNNs) and their variants have shown significant improvements and are gradually changing the narratives in weed identification and detection in farmlands[13], [15], [16].

Existing research highlights that Convolutional Neural Networks (CNNs) and their variants, especially You Only Look Once (YOLO) models, are highly effective for object detection [17], [18], [19], [20]. YOLO is particularly favoured over complex, slower Region-based Convolutional Neural Networks (R-CNN) methods due to its simpler structure, end-to-end training, and faster speed [15] which is crucial for real-time robotic applications, a point detailed in comparisons of CNN and YOLO architectures regarding structure, speed, and accuracy [13], [19], [21]. Furthermore, the YOLO algorithm stands out among single-stage object detection models for its robustness and efficiency, as it predicts bounding boxes and class probabilities directly from the image in a single pass, thereby achieving state-of-the-art speed and accuracy.

Consequently, this work leveraged the advances in YOLO variants and detection transformers to detect weeds and maize in a typical single-crop African farmland. The goal of this work is to conduct a comparative analysis of multiple YOLO architectures on the maize-weed dataset collected from multiple locations across the dry and wet seasons. This aims to determine which model yields optimal results, considering trade-offs among model size, accuracy, computational time, and speed. This study is an extension of an ongoing project to develop a robotic farm management system for African maize farmlands, aiming to create site-specific weed management (SSWM) using a lightweight Deep Learning (DL) approach, specifically focusing on developing an efficient ML/DL algorithm for the maize-weed environment in a traditional African setting.

II. MATERIALS AND METHODS

This work leverages the dataset acquired from (https://data.mendeley.com/datasets/jjbfcckrsp/2), as developed by Olaniyi et al. [22] , to investigate the performance of the varying single-stage object detection algorithms for the detection of maize-weed in a single-crop African Farmland system. This is part of a larger project geared towards developing a robust robotic solution for the management of maize-weed in single-crop African farmland.

In addition, this work investigated the models effects on spatio-temporal data acquired from different locations and climatic conditions. The methodology adopted in this work comprises the following stages: (1) Dataset acquisition, cleaning and pre-processing, (3) experimentation setup, (4) model development, investigation and evaluation, as depicted in Fig. 1.

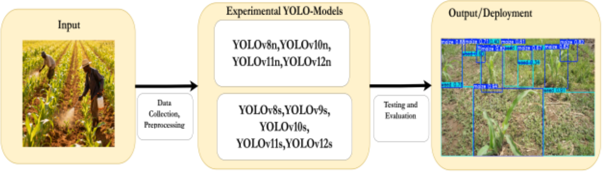


Fig.1. Methodology Description of the Study

A. Dataset Description

The dataset contains maize-weed images obtained from eighteen (18) different farmlands in Niger State, Nigeria. It contains images of maize-weed in situ in farmlands, obtained during both dry and wet seasons. The dataset provides the necessary spatio-temporal features for the investigation. However, 500 images from this dataset were annotated using the open-source tool LabelImg (https://github.com/HumanSignal/labelImg/releases). The tool applies bounding boxes to the images and saves them in the corresponding YOLO .txt format, which includes the coordinates (x, y, width, and height) and class names of the objects in the images. To ensure consistency and class balance during annotation, the numbers of weeds and maize were evenly labelled across the images in the dataset. Furthermore, the images used were a mix of images across the farmlands in different seasons. In this work, only two classes were defined: weed and maize. This is because the farming system practised on the farmlands was a single-cropping system, maize was the parent crop, while other plants were considered weeds. The entire dataset was divided into training, testing, and validation sets using a 70:15:15 split.

B. Model Description

This work focuses on the use of single-stage detector models (YOLO) for detecting maize-weed in typical Maize farmlands in Nigeria. The YOLO model was first introduced in 2015 by Redmon and his team at the University of Washington. [20], [23]. Moreover, since its introduction, the model has continued to attract attention, leading to numerous improvements to its initial architecture. YOLO is a single-stage object detector that uses a single CNN to directly predict bounding boxes and class probabilities across a grid, enabling real-time detection with reasonable accuracy. Across versions, the core architecture is a three-layer model: backbone, neck, and head. The backbone is for feature extraction; the neck, which serves as an intermediate layer between the backbone and the head, is for multi-scale fusion; and the head is for dense detection.

All versions of YOLOs retain these core layers for their operations [17], [20], [24], [25]. A detailed evolution map of the YOLO models is depicted in Fig. 2, showing the evolution and highlighting the remarkable improvement since its inception to date [20]. Irrespective of the version, YOLO relies on the Convolutional Neural Network (CNN) layers for its operation.

In the Backbone: CNN that converts the image into multi-scale feature maps; later versions use variants of Darknet, or more optimised convolution blocks. Neck: Feature pyramid that aggregates features from different scales to handle small, medium, and large objects. Head: Prediction layers that output bounding boxes, objectness scores, and class probabilities at each location, often with anchor-based or anchor-free designs. The model processes the image once and outputs all detections in a single forward pass, unlike two-stage detectors that use region proposals [20], [23].

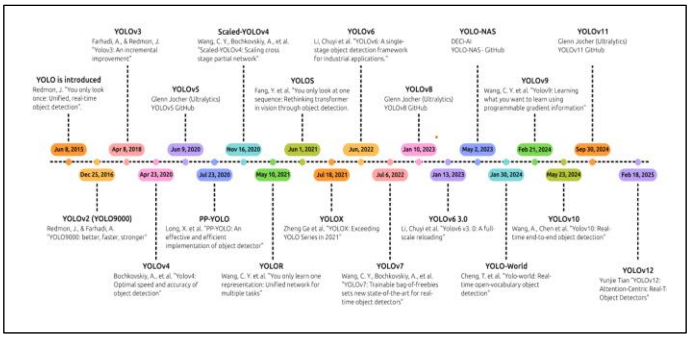


Fig. 2. Evolution of the YOLO Models [20].

In this work, the performance of YOLO versions 8, 9, 10, 11, and 12 was evaluated on a spatio-temporal maize-weed dataset, focusing only on the lightweight versions, nano (n) and small (s). This aligns with the project goal of applying lightweight models to detect maize weeds in farmlands. Parameters and model architecture are presented in Table 1.

TABLE I. DETAILS OF THE IMPLEMENTED YOLO MODELS

Model/Metrics	Input Size	Parameters (M)	FLOPs (Billion)	Model Size (MB)
YOLOv8n	640×640	3.2	8.7	6.5
YOLOv8s	640×640	11.2	28.6	22.5
YOLOv9s	640×640	7.2	26.7	26.7
YOLOv10n	640×640	2.3	6.7	4.7
YOLOv10s	640×640	7.2	21.6	15.1
YOLOv11n	640×640	2.6	6.5	5.4
YOLOv11s	640×640	9.4	21.5	19.1
YOLOv12n	640×640	2.6	6.5	50
YOLOv12s	640×640	9.3	21.4	115

C. Experimental Setup

In this study, nine (9) different experiments were conducted for each of the nine YOLO variants, namely YOLOv8n, YOLOv8s, YOLOv9s, YOLOv10n, YOLOv10s, YOLOv11n, YOLOv11s, YOLOv12n, and YOLOv12s. The training experiment was set up using Jupyter Notebook on a dedicated computer system with specifications defined in Table 2.

TABLE 2: EXPERIMENTAL SETUP ENVIRONMENT

Components	Specification
Environment	Jupyter Notebook
Computer	Lenovo LOQ
Processor	Intel Core i5
Specification	1.5 TB, 16GB RAM
GPU	NVIDIA GeForce RTX 5050, 8GB

For uniformity, fairness and consistency across the experiments, the models hyperparameters were kept the same across all experiments, and all models were obtained from Ultralytics (https://docs.ultralytics.com/models/). Table 3 lists the model hyperparameters used across all experiments.

TABLE 3: MODEL HYPERPARAMETERS

Parameters	Values
Device	GPU
Image size	640x640
Epochs	100
Optimiser	AdamW
Learning Rate	0.01
Batch	16
Workers	8

D. Model Evaluation

Model evaluation is based on mean Average Precision (mAP), which provides an overall measure of accuracy by averaging the precision-recall curve across all object classes. It also measures the proportion of correct positive predictions among all positive predictions, helping to minimise false positives. The Recall measures the proportion of actual positive cases correctly identified, ensuring that all relevant objects are detected.

Another metric adopted is Intersection over Union (IoU), which measures the overlap between the predicted and ground-truth bounding boxes and is used to fine-tune localisation accuracy. Additionally, inference speed was evaluated to quantify how quickly the model processes an image, typically expressed in frames per second (FPS), a metric critical for real-time applications. These metrics are defined as follows:

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

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

$$mAP = \frac{1}{n}\sum_{I=1}^{n} AP_i \qquad (3)$$

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall} \qquad (4)$$

III. RESULTS AND DISCUSSION

The study investigated the performance of YOLO model variants suitable for deployment in lightweight applications, based on size, parameters, and their ability to detect and differentiate weeds from the parent crop under review. The YOLO models considered are the small and nano variants as defined in Table 1. These models were trained using the annotated maize-weed dataset obtained from the repository, as described in the data description. Furthermore, to ensure consistency, fairness and uniformity in performance and evaluation, all models were subjected to the same hyperparameters and experimental setup, as described in Table 3.

The model outputs during training are shown in Fig. 3a and 3b, with 3a depicting the performance of the small (s) models and 3b that of the nano (n) models. The performance results of the models with respect to accuracy and F1-Score are depicted in Fig. 3a and 3b, showing how performance changes as the epoch count increases. Across all models in the nano and small categories, accuracy increases with the number of epochs, indicating that learning occurs. This supports the fact that deep learning model performance improves with increasing training epochs [18].

Furthermore, this pattern is observed in our training results, as depicted in Fig. 3a. As the training epoch increases, performance improves; however, as the training epoch approaches 100, the model converges to a relatively stable point after a period of fluctuation. This phenomenon confirms that the models are learning; accordingly, an epoch of 100 was used in this study. Notably, YOLOv10n and YOLOv10s achieve the lowest performance, whereas YOLOv8n and YOLOv8s achieve the highest, slightly above the other models.

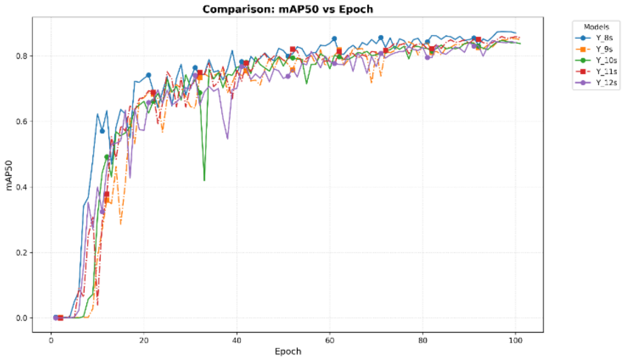


Fig 3a: Comparative Map@50 Performance of the Small Models with Epoch

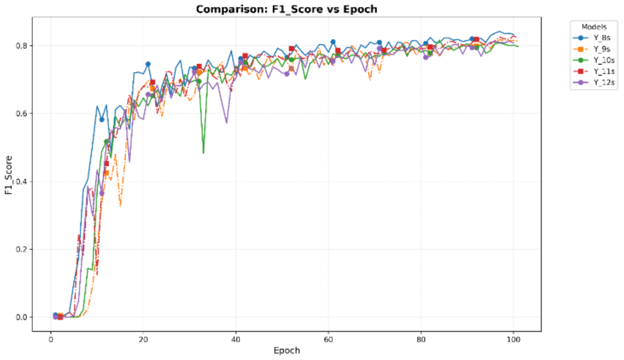


Fig 3b: Comparative Map@50 Performance of the Nano Models with Epoch

Comparatively, the training time was investigated, and the results are presented in Fig. 4. The results show that as computational time increases, the model parameters and size increase. Observe that the nano variants have lower computational costs than their small variants; however, YOLOv9s has a considerably higher computational cost than all models. The complexity, model sizes, and parameters of the models contributed to the trade-off in computational time; however, other factors, such as dataset size, may have accounted for the variation in training time.

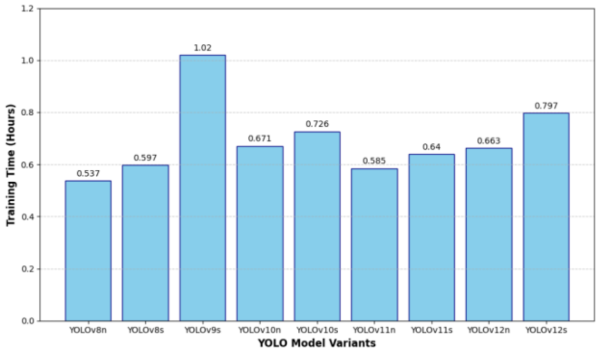


Fig. 4. Comparative Performance of the Model with Training Time

A. *Comparative Performance of the Models*

The performance of the YOLO variants considered in this study was evaluated on the specified datasets and hyperparameters, as presented in Table 4, which reports Precision, Recall, F1 Score, mAP@50, mAP@50-95, and Inference Time for the models for both classes (maize and weeds). The results show model precision ranging from 0.779 to 0.853, with YOLOv10n and YOLOv10s achieving the highest and lowest performance, respectively. Furthermore, recall ranges from 0.805 to 0.866, with YOLOv10n and YOLOv8s achieving the highest and lowest performance, respectively. Additionally, the F1 scores have a range from 0.806 to 0.832, with the YOLOv8s and YOLOv10n achieving the best and least performance, while the mAP@50 values range from 0.844 to 0.870, with YOLOv8s offering the best performance with a value of 0.870 and YOLOv12s having the least performance with a value of 0.844.

TABLE 4: COMPARATIVE PERFORMANCE ANALYSIS OF THE MODELS

Models	Precision	Recall	F1 Score	mAP @50	mAP @50-95	GFLOP
YOLOv8n	0.825	0.825	0.825	0.869	0.626	8.1
YOLOv8s	0.801	**0.866**	**0.832**	**0.870**	**0.637**	28.4
YOLOv9s	0.812	0.818	0.815	0.851	0.605	26.7
YOLOv10n	**0.852**	0.765	0.806	0.840	0.593	6.5
YOLOv10s	0.779	0.837	0.807	0.846	0.606	21.4
YOLOv11n	0.816	0.805	0.810	0.862	0.609	6.3
YOLOv11s	0.828	0.829	0.828	0.858	0.615	21.3
YOLOv12n	0.823	0.806	0.814	0.860	0.604	6.3
YOLOv12s	0.798	0.833	0.815	0.844	0.585	21.2

Furthermore, in the same pattern, YOLOv8s achieves the best mAP@50-95, whereas YOLOv12s achieves the lowest. However, with respect to the inference time shown in Fig. 5, the models vary from 4.4ms to 12.3ms per frame. YOLOv8n had the fastest inference time of 4.4ms, while YOLOv12s had the slowest inference time. The models' inference performance is shown in Fig. 6.

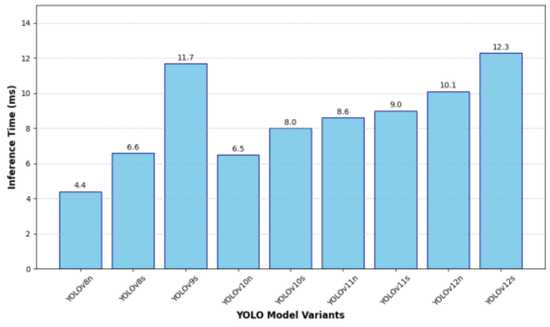
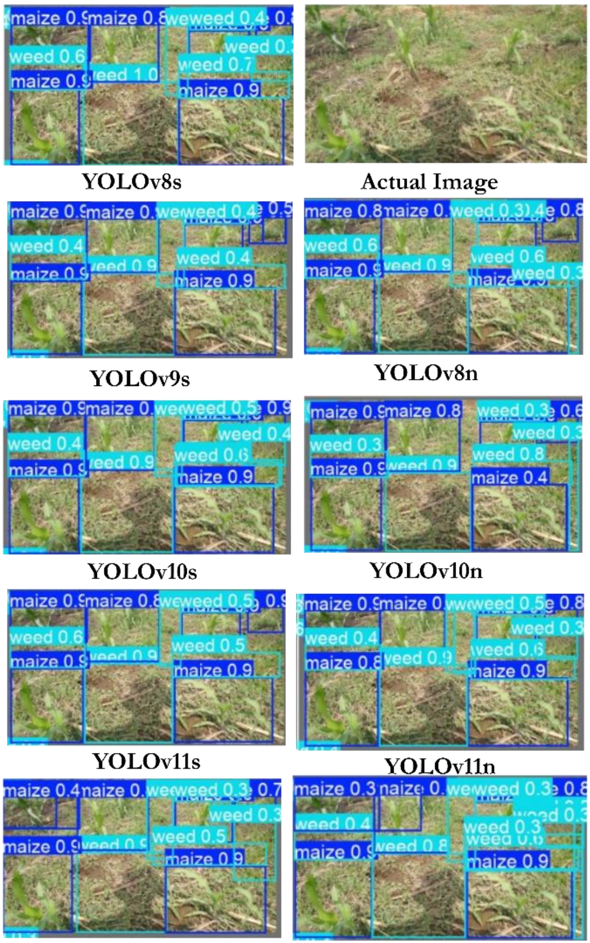


Fig. 5. Comparative Performance of the Model with Inference Time



YOLOv8s Actual Image

YOLOv9s YOLOv8n

YOLOv10s YOLOv10n

YOLOv11s YOLOv11n

YOLOv12s YOLOv12n

Fig. 6. Inference performance of the models

B. Discussion of Results

This study investigates the performance of the nine YOLO model variants for detecting maize-weed in a typical maize farmland. The dataset comprises various weed species and maize crops at different growth stages. The model's performance on the dataset is presented in Table 4. The progression from YOLOv8 to YOLOv12 shows a shift towards efficiency, in terms of computational overhead, the Small (s) models consistently demand more resources, with YOLOv8s requiring the most at 28.4 GFLOPs. However, architectural refinements in the newer YOLOv12s have reduced this to 21.2 GFLOPs despite an increase in layer depth (159 layers), suggesting a more optimised use of feature-extraction blocks. The Nano (n) models remain the most efficient, with YOLOv12n and YOLOv11n operating at 6.3 GFLOPs. This efficiency is reflected in the training times. YOLOv8n converged the fastest at 0.537 hours, whereas YOLOv9s required a significantly longer 1.022 hours. This suggests that the Programmable Gradient Information (PGI) used in the v9 architecture, while innovative, increases the computational burden during backpropagation relative to the more streamlined v10, v11, and v12 iterations.

Furthermore, the models exhibit distinct characteristics with respect to Precision (P) and Recall (R). YOLOv10n achieved the highest precision of 0.853, which is critical for reducing false-positive detections that could lead to the accidental destruction of crops. Conversely, YOLOv8s achieved the highest recall (0.861), making it the most reliable option for ensuring comprehensive weed removal, as it is less likely to miss invasive plants hidden in the canopy.

In addition, the mean Average Precision mAP@50 results show that the older YOLOv8s still maintains a slight lead at 0.870, followed closely by YOLOv8n at 0.869. The newer models, such as YOLOv11n (0.861) and YOLOv12n (0.860), provide nearly identical performance while weighing significantly less (5.5 MB compared with 22.5 MB for v8s), indicating that modern YOLO versions have achieved a higher accuracy-per-parameter ratio.

The deeper mAP50-95 metric, which measures localisation accuracy across multiple Intersection over Union (IoU) thresholds, highlights the robustness of the v8 family, with YOLOv8s achieving a peak of 0.638. However, the newer variants, such as YOLOv12n (0.604) and YOLOv11s (0.615), show slight regression on this metric, suggesting that although they are faster, they may be slightly less precise in delineating bounding boxes in dense agricultural scenes. This is likely an inference from the dataset size: larger models (Smaller variants) generally struggle with overfitting on smaller datasets, whereas the nano variants remain more stable. Across all nine models, a consistent performance gap is observed between maize detection (approximately 0.94 mAP@50) and weed detection (ranging from 0.748 to 0.805 mAP@50), underscoring that the morphological diversity of weeds remains a significant challenge despite advances in architecture.

In the context of the problem at hand, rapid detection of weed-crops necessitates that inference speed be the primary metric for real-time field deployment. The YOLOv8n delivered the fastest raw inference at 4.4ms, while the small variants ranged from 6.6ms (YOLOv8s) to 12.3ms (YOLOv12s). These results indicate that the nano models, particularly YOLOv8n and YOLOv12n, are best suited for high-speed detection applications, while for slow-speed detection, YOLOv8s remains the superior choice for maximised weed detection due to its higher recall and mAP@50-95 scores.

IV. CONCLUSION

The application of deep learning techniques for weed detection in farmland has proven effective and efficient compared to traditional approaches. This study investigated the performance of YOLO model variants suitable for deployment in lightweight applications, based on model size and parameter count, and their ability to detect and distinguish weeds from the parent crop (Maize). This study investigated the performance of nine YOLO model variants, with a focus on small- and nanosized models. This aims to develop a lightweight model for integration into a real-time robotic system to detect and manage weeds in typical Maize farmlands.

A comprehensive evaluation of the YOLO architecture from YOLOv8 to YOLOv12 reveals a distinct trade-off between computational efficiency and detection robustness tailored to precision agriculture. The analysis indicates that deployment platform constraints must split the model selection process. For high-speed detection applications such as aerial detections, YOLOv8n is the primary recommendation, with an inference latency of 4.4ms and a high mAP@50 of 0.869. Conversely, for applications in which high recall is vital for complete weed eradication, YOLOv8s remains superior, achieving a recall of 0.861.

However, for a more robust, lighter application, such as edge-based mobile devices, YOLOv12n strikes the best balance, offering modern architectural optimisations in a compact 5.5MB footprint and 6.3 GFLOPs. Notably, the YOLOv8 family maintains strong raw metrics on this dataset; the reduced model size and lower GFLOPs of the YOLOv12 series indicate a transition toward more energy-efficient edge computing. The results obtained from the YOLO models demonstrate their capability to serve as a lightweight application for detecting weeds in conventional farmland, thereby supporting the adoption of precision agriculture to increase productivity and efficiency.

ACKNOWLEDGMENT

This research was supported by the Artificial Intelligence Exploration Center (AIX) of SUNY Polytechnic Institute, Utica, NY, USA, Africa Centre of Excellence for Mycotoxin and Food Safety (ACEMFS), Federal University of Technology, Minna and the National Universities Commission (NUC) of Nigeria.

REFERENCES

[1] P. Biswas and S. K. Goswami, “Precision Farming Utilizing Internet of Things, Artificial Intelligence and Automation: An Overview,” *International Journal of Engineering and Information Management*, vol. 01, no. 01, pp. 60–72, Jan. 2025, doi: 10.52756/ijeim.2025.v01.i01.005.

[2] A. Ehrampoosh *et al.*, “Intelligent weed management using aerial image processing and precision herbicide spraying: An overview,” Aug. 01, 2025, *Elsevier Ltd*. doi: 10.1016/j.cropro.2025.107206.

[3] A. Bodhale and S. Verma, “Review of Segmentation Techniques for Weed Detection in Agricultural Crops,” *Procedia Comput Sci*, vol. 259, pp. 61–70, 2025, doi: 10.1016/j.procs.2025.03.307.

[4] M. Vasileiou *et al.*, “Transforming weed management in sustainable agriculture with artificial intelligence: A systematic literature review towards weed identification and deep learning,” *Crop Protection*, vol. 176, Feb. 2024, doi: 10.1016/j.cropro.2023.106522.

[5] A. Gómez, H. Moreno, C. Valero, and D. Andújar, “Spatio-temporal stability of intelligent modeling for weed detection in tomato fields,” *Agric Syst*, vol. 228, Aug. 2025, doi: 10.1016/j.agsy.2025.104394.

[6] W. Luo *et al.*, “Real-time identification and spatial distribution mapping of weeds through unmanned aerial vehicle (UAV) remote sensing,” *European Journal of Agronomy*, vol. 169, Aug. 2025, doi: 10.1016/j.eja.2025.127699.

[7] H. R. Qu and W. H. Su, “Deep Learning-Based Weed–Crop Recognition for Smart Agricultural Equipment: A Review,” Feb. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/agronomy14020363.

[8] R. Raja, W. H. Su, D. C. Slaughter, and S. A. Fennimore, “Real-time precision crop identification in high weed-density environments for robotic weed control using spectral fluorescence imaging in celery,” *Comput Electron Agric*, vol. 231, Apr. 2025, doi: 10.1016/j.compag.2025.110022.

[9] A. C. Buzanini, R. H. Furlanetto, A. W. Schumann, and N. S. Boyd, “Integration of targeted herbicide application technologies into an integrated weed management program for tomato plasticulture production,” *Smart Agricultural Technology*, vol. 11, Aug. 2025, doi: 10.1016/j.atech.2025.100981.

[10] S. Meesaragandla, M. P. Jagtap, N. Khatri, H. Madan, and A. A. Vadduri, “Herbicide spraying and weed identification using drone technology in modern farms: A comprehensive review,” *Results in Engineering*, vol. 21, Mar. 2024, doi: 10.1016/j.rineng.2024.101870.

[11] A. Upadhyay, S. G C, Y. Zhang, C. Koparan, and X. Sun, “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*, vol. 18, Dec. 2024, doi: 10.1016/j.jafr.2024.101331.

[12] G. C. Sunil, A. Upadhyay, and X. Sun, “Development of software interface for AI-driven weed control in robotic vehicles, with time-based evaluation in indoor and field settings,” *Smart Agricultural Technology*, vol. 9, Dec. 2024, doi: 10.1016/j.atech.2024.100678.

[13] M. G. Rahman *et al.*, “ADeepWeeD: An adaptive deep learning framework for weed species classification,” *Artificial Intelligence in Agriculture*, vol. 15, no. 4, pp. 590–609, Dec. 2025, doi: 10.1016/j.aiia.2025.04.009.

[14] S. Nigam, A. Dheeraj, H. Sachan, and S. Marwaha, “Automated weed classification using attention-embedded ConvNeXtV2 architecture,” *Procedia Comput Sci*, vol. 260, pp. 291–299, 2025, doi: 10.1016/j.procs.2025.03.204.

[15] C. Russo, V. Cirillo, M. Esposito, M. Lentini, N. Pollaro, and A. Maggio, “Convolutional neural network for the early identification of weeds: A technological support to biodiversity and yield losses mitigation,” *Smart Agricultural Technology*, vol. 9, Dec. 2024, doi: 10.1016/j.atech.2024.100594.

[16] G. C. Sunil *et al.*, “Field-based multispecies weed and crop detection using ground robots and advanced YOLO models: A data and model-centric approach,” *Smart Agricultural Technology*, vol. 9, Dec. 2024, doi: 10.1016/j.atech.2024.100538.

[17] A. Tripathi, M. K. Gupta, C. Srivastava, P. Dixit, and S. K. Pandey, “Object Detection using YOLO: A Survey,” in *Proceedings of 5th International Conference on Contemporary Computing and Informatics, IC3I 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 747–752. doi: 10.1109/IC3I56241.2022.10073281.

[18] O. G. Ajayi, J. Ashi, and B. Guda, “Performance evaluation of YOLO v5 model for automatic crop and weed classification on UAV images,” *Smart Agricultural Technology*, vol. 5, Oct. 2023, doi: 10.1016/j.atech.2023.100231.

[19] J. Ju, X. Ma, C. Yang, F. Qian, J. Wang, and G. Chen, “Lightweight detection method for weeds in complex paddy fields based on the improved YOLOv5,” *Smart Agricultural Technology*, p. 101418, Sep. 2025, doi: 10.1016/j.atech.2025.101418.

[20] N. Jegham, C. Y. Koh, M. Abdelatti, and A. Hendawi, “YOLO Evolution: A Comprehensive Benchmark and Architectural Review of YOLOv12, YOLO11, and Their Previous Versions,” Mar. 2025, [Online]. Available: http://arxiv.org/abs/2411.00201

[21] A. Upadhyay, S. G C, S. Das, J. Mettler, K. Howatt, and X. Sun, “Multiclass weed and crop detection using optimized YOLO models on edge devices,” *J Agric Food Res*, vol. 22, Aug. 2025, doi: 10.1016/j.jafr.2025.102144.

[22] O. M. Olaniyi *et al.*, “Development of maize plant dataset for intelligent recognition and weed control,” *Data Brief*, vol. 47, p. 109030, Apr. 2023, doi: 10.1016/J.DIB.2023.109030.

[23] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *In Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 779–788. [Online]. Available: http://pjreddie.com/yolo/

[24] C. Liu, Y. Tao, J. Liang, K. Li, and Y. Chen, “Object Detection Based on YOLO Network,” in *Proceedings of 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC 2018) : December 14-16, 2018, Chongqing, China*, IEEE Press, 2018, pp. 799–803.

[25] Y. Chen, M. C. Goorden, F. J. Beekman, and R. Cheng, “A survey: Comparison between Convolutional Neural Network and YOLO in image identification,” *Journal of Physics: Conference Series, 2020 iopscience.iop.org*, p. 12139, 2020, doi: 10.1088/1742-6596/1453/1/012139/META.