International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I10P104

Volume 72 Issue 10, 35-45, October 2024

Original Article

Accurate and Efficient Real-Time Crop and Weed Identification Using YOLO-NAS: A Neural Architecture Search-Optimized Object Detection Model

Sanskruti Patel[1], Dharmendra Patel[2], Atul Patel[3], Jay Nanavati[4], Unnati Patel[5], Ankit Faldu[6], Amisha Shingala[7]

[1,2,3,4,5,6,7]*Smt. Chandaben Mohanbhai Patel Institute of Computer Applications, Charotar University of Science and Technology Changa, Gujarat, India.*

[1]*Corresponding Author : sanskrutipatel.mca@charusat.ac.in*

Received: 20 May 2024 Revised: 27 September 2024 Accepted: 02 October 2024 Published: 25 October 2024

Abstract *- Accurate identification of plants and weeds is essential for precision farming since it allows for focused treatments, effective use of resources, and reduced environmental impact. In order to improve its architecture for crop and weed detection tasks, this research proposed a YOLO-NAS, an effective object detection model that makes use of Neural Architecture Search (NAS). With the use of cutting-edge methods like selective quantization, mixed precision training, and knowledge distillation, YOLO-NAS may be deployed seamlessly across a range of computing resources. A large dataset of agricultural images was used to assess the performance of the model while considering evaluation parameters, including mean Average Precision (mAP), recall, and precision. The results of the experiments show that YOLO-NAS works better than most object detection models, with a mAP of 86.11% and a balance between weed misdetection minimization and crop identification accuracy. With its high accuracy, real-time functionality, and easy deployment, the suggested model is a viable option for automated crop and weed identification that helps in precision agriculture.*

Keywords *- Computer vision, Crop and weed object detection, Deep learning, YOLO, YOLO-NAS.*

1. Introduction

Precision agriculture has emerged as a crucial approach to address the growing global demand for food while minimizing environmental impact and optimizing resource utilization. The capacity to recognize and discriminate between weeds and crops in agricultural areas with accuracy is one of the significant components of precision agriculture. The manual identification of crops and weeds used in traditional weed management methods can be tedious, costly, and time-consuming [1]. For a number of reasons, accurate crop and weed identification is essential to precision agriculture. One of the most important factors is targeted weed control, which calls for precision herbicide application or manual removal of weeds based on their correct identification and localization. This strategy optimizes resource efficiency while minimizing the use of herbicides and their adverse impacts on the environment. Accurately identifying and differentiating crop plants from weeds allows farmers to monitor crop development, evaluate crop health, and increase yield estimates [2]. Accurate crop and weed detection is essential for the development of robotic and autonomous systems for tasks like weeding and harvesting. Accurate identification and precise location of these plants allow autonomous systems to function effectively and efficiently in farming. Making decisions based on data is made easier by accurate crop and weed identification. It produces useful data that decision support systems and platforms for precision agriculture can incorporate. With the help of this data-driven strategy, farmers may increase production and profitability by making well-informed decisions based on current field conditions [3]. Recent advancements in deep learning, a subfield of artificial intelligence, have opened exciting possibilities for automated crop and weed detection. Deep learning algorithms can learn complex patterns from large datasets of images that enable them to recognize objects with remarkable accuracy. In deep learning, object detection techniques are proven to excel at identifying and locating specific objects within an image. These methods, which include SSD (Single Shot MultiBox Detector), YOLO (You Only Look Once), and Faster RCNN (Convolutional Neural Networks), have shown excellent performance across a range of applications [4]. However, when applied to the agricultural domain, current object detection models face difficulties. These challenges include issues with real-time performance, flexibility in deployment, and the capacity to identify weeds from crops in a variety of field circumstances. This study suggests a novel YOLO-NAS object detection model for precise and effective real-time crop and weed identification in order to overcome these issues. YOLO-NAS leverages Neural Architecture Search (NAS) techniques to optimize its architecture. YOLO with NAS integration enables superior performance and adaptability to the complex requirements of agricultural object detection [5]. The proposed YOLO-NAS achieves high accuracy and efficiency in identifying and localizing both crops and weeds. With the use of advanced training methods, YOLO-NAS may be deployed with ease on a variety of computing platforms like robust workstations, resource-constrained edge devices, etc. The proposed YOLO-NAS model represents a significant advancement in the field of detecting crops and weeds and offering a robust and efficient solution to enhance precision agriculture practices.

The major research contributions of this paper are given below;

- Proposing YOLO-NAS, a competent and accurate object detection model for crop and weed detection tasks, leveraging Neural Architecture Search (NAS) to optimize its building blocks.
- Demonstrating YOLO-NAS performance compared to state-of-the-art object detection models, achieving a mAP of 86.11% and a prominent balance between accurately identifying crops and minimizing weed misdetection.
- Presenting a quantization-friendly and deployment-flexible solution that can be seamlessly integrated into various agricultural operations.
- Providing a detailed architecture and implementation guide for YOLO-NAS, facilitating further research in the area of automated crop and weed detection for precision agriculture.

2. Related Work

Ye. Mu et al. (2022) present a model that is based on Faster R-CNN for identifying weed seedlings in agricultural fields. The model achieves high accuracy (95.61%), recall (87.26%), F1-score (91.24%), and mean intersection over union (93.7%) in detecting various weed species, including charlock, maize, scentless mayweed, fat hen, common wheat, sugar beet, and shepherd's purse [6]. The proposed model outperforms the standard Faster R-CNN. The authors highlight the potential of their approach for improving the management of weeds and reducing the use of herbicides. N.Y. Murad et al. (2023) presented a systematic literature review on weed detection using deep learning, revealing varying performance levels across algorithms, with some demonstrating high accuracy [7]. M. H. Saleem et al. (2022) focus on enhancing weed detection using a Faster RCNN ResNet-101 model with an optimized approach for anchor boxes.

The study achieved a high accuracy rate of 96.02% in weed detection [8]. The research is published under an open-access license, allowing for widespread distribution and reference. The methodology involves improving the anchor box approach to enhance the model's performance in detecting various types of weeds. The FT_BRC picture collection, presented by V. N. Thanh Le et al. (2021), consists of 3,380 photographs taken on a commercial farm in the Western part of Australia. They have used real-world field settings using a camera mounted on a movable trolley [9]. A subset of this dataset was thoroughly annotated in order to locate targeted weeds and estimate weed density. Various feature extractors are combined and used with Faster RCNN models for detecting weed presented in the field. The network outperformed other networks, as evidenced by the results obtained during the experiment, which show that it acquired a mean average accuracy (mAP) of 0.555. H. Peng et al. (2022) propose an algorithm for the exposure of weeds in paddy fields [10]. They established a dataset containing both rice and weeds to train and evaluate their model. Additionally, the team introduced a method for localizing rice crops within the images. The proposed model demonstrated impressive results, achieving an accuracy of 94.1% and a processing speed of 24.3 frames per second. According to J. M. L. Correa et al. (2021), a one-step method for weed detection and classification was created utilizing the popular Object Detection Network called RetinaNet [11]. The system focused on identifying different weeds with two growth stages in the field of maize crop. The accuracy of the predictions was assessed using the mean Average Precision (mAP) metric, resulting in a score of 0.88, ranging from 0.98 to 0.75 across different classes. K. N. Sudheer et al. (2023) propose an algorithm that identifies weeds and crops based on predefined plant characteristics, locating weeds in each row and measuring the distance between them [12]. The input data consists of 4-channel NIR + RGB or regular RGB images, depending on the sensor used. By using a large dataset of cotton weeds and crops to train a model that is applicable to other crops, their goal was to increase the accuracy of weed detection. The SSD Mobilenet model achieves 90-95% accuracy in this context O. G. Ajayi (2023) presents a study that evaluated the YOLOv5 model for classifying crops and weeds on UAV images [13]. It categorized different crops from weeds, training the model across epochs to optimize performance. J. Chen et al. (2022) present the model based on YOLOv4, which enhances sesame and weed detection by incorporating local importance pooling for attention, SE blocks for logic improvement, and ASFF structure to address detection gaps [14].

This model achieves high accuracy while maintaining fast detection speed. V. S. Babu and N. Venkatram (2024) present YOLOv4 that excels in weed detection and localization in soybean fields, achieving 98.42% accuracy, 93.13% recall, and 81.24% mAP, outperforming R-CNN and SSD networks [15]. Various pre-trained models are explored for weed/crop classification, with Densenet201 showing the highest accuracy at 99.67%. YOLOv4 demonstrates efficiency in both classification and detection with localization capabilities. Table 1 summarizes the different object detection models with their advantages, disadvantages and applications.

3. Research Gap

While existing deep learning models like Faster R-CNN, SSD, YOLO, and RetinaNet offer promising solutions for crop and weed detection, some limitations require further research as follows;

- Investigating techniques for real-time object detection with high accuracy, potentially using hardware acceleration.
- Develop larger and more diverse datasets encompassing various crops, weeds, and agricultural environments.
- The cost-effectiveness of deploying these models, especially with complex hardware requirements, needs further investigation.

4. A Proposed Method for Crop and Weed Detection using YOLO with Neural Network Search (YOLO-NAS)

In this research, an efficient architecture for object detection called YOLO-NAS is proposed for crop and weed detection. YOLO-NAS builds upon the conventional YOLO framework by leveraging Neural Architecture Search (NAS) to discover an optimal architecture for object detection tasks automatically. Deci AI released YOLO-NAS, a state-of-the-art object identification model that comes in three flavors: small, medium, and large [16]. In this study, YOLO-NAS is proposed to address the challenging task of crop and weed recognition in agricultural fields.

YOLO-NAS uses Neural Architecture Search (NAS) technology to improve its architecture on its own, making it more suitable for production-level performance and improving its real-time object identification capabilities. The quantization-friendly fundamental block architecture of YOLO-NAS, which overcomes earlier drawbacks in YOLO models, is a noteworthy feature. This quantization based design enables its deployment on resource-constrained devices like smartphones and IoT devices [17]. This feature is particularly beneficial in agricultural settings, where real-time object detection on edge devices can facilitate timely and efficient decision-making.

YOLO-NAS incorporates advanced training methods with post-training quantization techniques [18]. These techniques enhance the robustness and efficiency of the model, ensuring reliable performance in diverse field conditions. Neural Architecture Search (NAS) in our research was employed to optimize the YOLO-NAS model for detecting crops among weeds. The search space encompassed key architectural elements such as convolutional layers, kernel sizes, activation functions, and skip connections. For the search strategy, we utilized an evolutionary algorithm that iteratively refined candidate architectures based on performance, combining and mutating high-performing models over multiple generations. To ensure computational efficiency, we adopted a weight-sharing approach in the performance estimation strategy, allowing multiple architectures to share parameters during training.

A detailed explanation of building blocks of YOLO-NAS is given below:

1. Input Image: The model takes an image as input, typically with three channels representing Red, Green, and Blue (RGB) color information. The image dimensions are denoted as W (width), H (height), and 3 channels (C).

2. Early Feature Extraction: A convolutional layer performs an initial feature extraction from the input image. This layer extracts low-level features like edges and textures. The output is a feature map with potentially reduced dimensions (W', H') and an increased number of channels (C') compared to the input image.

3. Dual-Path Backbone: This is the core of the YOLO-NAS architecture where NAS is implemented:

- Dense Path: This path consists of multiple Dense Blocks. Each dense block is a group of convolutional layers with skip connections. Skip connections allow information to flow directly from earlier layers to later ones while preserving detailed features crucial for object detection. It is quite efficient especially for smaller objects like detecting weeds.

- Sparse Path: This path utilizes Transition Blocks to reduce the spatial resolution (W" x H") of the feature maps while increasing the number of channels (C"). This helps capture higher-level semantic information about the image content. This is beneficial for detecting larger objects like crops.

4. Feature Concatenation: Outputs from both Dense and Sparse paths are concatenated. This combines the detailed features from the dense path with the semantic information from the sparse path. It creates a richer feature representation suitable for detecting objects of various sizes (crops and weeds).

5. Neck: The neck further refines the concatenated features from the backbone. It might use operations like upsampling and concatenation to create a feature pyramid with different resolutions. It allows the model to discover objects of various sizes effectively.

6. Object Detection: The processed features from the neck are fed into the head of the network for object detection tasks. It includes:

- Classification Branch: Predicts the class probabilities for each detected object (e.g., crop or weed).
- Regression Branch: Predicts bounding box coordinates for each detected object.

Table 1. Summary of object detection models used for crop and weed detection

Model	Advantages	Disadvantages	Applications	References
Faster R-CNN	High accuracy; Good at handling complex backgrounds	Slower than SSD and YOLO; Requires large training datasets.	Precise crop and weed detection, even in cluttered fields.	[6], [7], [9]
SSD (Single Shot MultiBox Detector)	Faster than Faster R-CNN; Good real-time performance	Lower accuracy compared to Faster R-CNN.	Weed detection in fields captured by drones or ground vehicles.	[7], [13]
YOLO (You Only Look Once)	Very fast processing speed; Efficient for resource-constrained devices	Lower accuracy than Faster R-CNN for complex scenes.	Real-time crop and weed detection on mobile platforms.	[7], [14], [13], [15]
RetinaNet	High accuracy comparable to Faster R-CNN; Faster inference speed than Faster R-CNN	More complex architecture compared to SSD and YOLO. It may require larger training datasets.	Crop and weed detection in diverse agricultural environments.	[10], [11], [7]

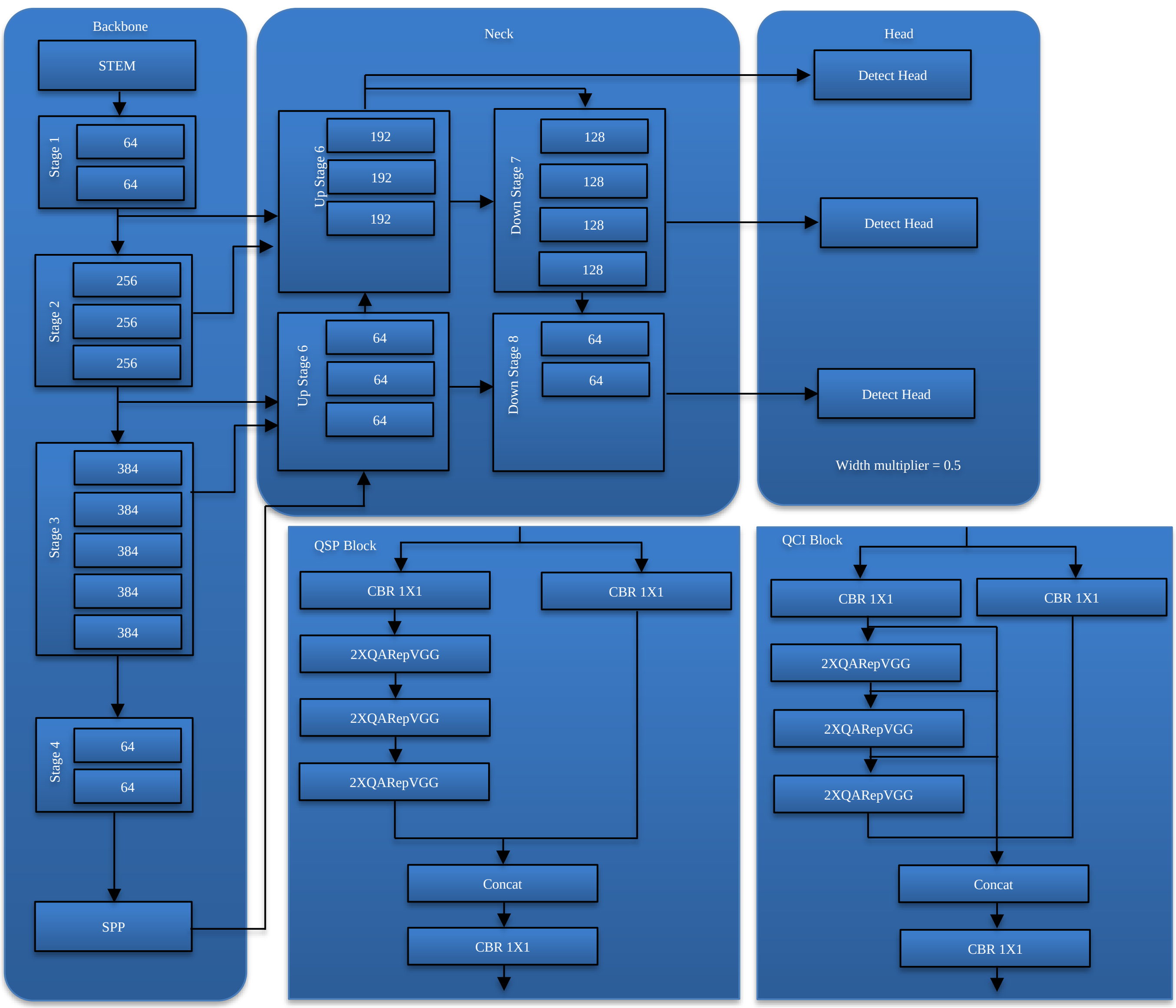

Fig. 1 A detailed architecture of YOLO-NAS for crop and weed detection [16]

A supervised learning approach is followed for training the proposed YOLO-NAS for crop and weed detection. Pseudo-labels were produced with the COCO dataset following the model's pre-training on the two million picture and 365 category Objects365 [20] dataset.

Finally, the models are trained using the original 118,000 training images from the COCO dataset. The model was learned from labeled images of agricultural fields. Each image is tagged with bounding boxes and class information for crops and weeds.

During training, the model minimized a loss function, combining bounding box and classification losses and refining its predictions to align with the ground truth [21]. After training, YOLO-NAS demonstrated remarkable adaptability and deployment across multiple platforms, such as CPUs, GPUs, and TPUs.

Additionally, quantization techniques further optimized the size and performance of the model on hardware with limited computational resources. It is enabling its practical application in resource-constrained agricultural settings. Algorithm 1 outlines the specific steps used to perform object detection tasks in this research.

An Algorithm for Crop and Weed Detection

1. *Input: Dataset D = {(I_i, B_i, C_i)} where:*
   I_i is an agricultural field image.
   B_i is a set of bounding boxes for objects (crops and weeds) in image I_i.
   C_i is a set of class labels (crop or weed) corresponding to bounding boxes in B_i.
2. *For each image I_i in dataset D:*
   Resize I_i to a fixed size (H, W, C) suitable for the model.
   Normalize pixel values in I_i to a range (e.g., 0-1)
3. *Divide the preprocessed data* `D` *into training set* `D_train`*, validation set* `D_val`*, and test set* `D_test` *using a predefined split ratio.*
4. *Initialize the YOLO-NAS model architecture with parameters* `θ`*:*
5. *Define Dense path as* `f_d`*; Sparse paths using* `f_s` *; Neck as* `f_neck`*; classification as* `f_cls`*; bounding box regression as* `f_reg`*;*
   Load pre-trained weights `θ_pre` *for the YOLO-NAS model;*
6. *Set hyperparameters for training;*
   Optimizer function `opt = Adam;` *Learning rate* `η = 0.0001; Number of training epochs N = 30; Batch size b=32;`
7. *Define the loss function* `L` *as a combination of bounding box loss* `L_box` *and* `classification loss L_cls:`
   `L(θ, I, B, C) = L_box(f_reg(f_neck(f_d(I, θ_d) + f_s(I, θ_s))), B) + λ * L_cls(f_cls(f_neck(f_d(I, θ_d) + f_s(I, θ_s))), C)` *where λ is a hyperparameter balancing the importance of each loss term.*
8. *For each epoch n in [1, N]:*
9. *For each batch B_t of images and corresponding labels from D_train:*
   Forward pass - Predict bounding boxes and class probabilities: `ŷ_b = f_reg(f_neck(f_d(I_j, θ_d) + f_s(I_j, θ_s)))` `ŷ_c = f_cls(f_neck(f_d(I_j, θ_d) + f_s(I_j, θ_s)))` *for all images* `I_j` *in* `B_t`*.*
   Loss calculation - L_batch = L(θ, B_t, B_truth, C_truth) where B_truth and C_truth are ground truth bounding boxes and class labels in B_t.
   Backpropagation - ∇_θ L_batch to compute gradients for each model weight in θ.
   Weight update - θ ← opt(θ, η, ∇_θ L_batch) using the optimizer function with learning rate.
10. *Evaluate model performance on* `D_val` *using metrics like mean Average Precision (mAP). Save the trained model weights* `θ*`*.*

5. Experiments and Results

5.1. Dataset

In this experiment, the annotated food crops and weed photos dataset was used. The Life Sciences and Technologies, Latvia University and the Institute of Electronics and Computer Science, Latvia, made this dataset available in 2021 [22]. A dataset that includes hand-annotated copies of images showing food crops and weeds at various stages of seedling growth. 1118 photos (.jpg files) and 7853 XML annotations (.xml files) make up the dataset. It includes 14 staple food crops and common weed species that grow in Latvia's agricultural fields.

The dataset was developed to support the creation and comparison of deep learning and computer vision algorithms for precision agriculture, particularly for applications like robotic weed control management, crop recognition, and weed detection. The dataset contains objects belonging to two classes: weed and crop. All images in the dataset have bounding box annotations for the labeled objects.

This dataset seems to be a valuable resource for developing and evaluating object detection models tailored for agricultural applications, particularly in identifying and distinguishing between weeds and crops during their early growth stages. Figure 2 shows the raw images of food crops and weeds from the dataset. Figure 3 shows the colored bounding boxes on the annotated image of the weed and crop.

Fig. 2 Raw images of food crops and weeds of the dataset

Fig. 3 Colored bounding boxes on annotated images of the weed and crop

5.2. Hardware and Software Setup

The experiments leveraged a powerful computational environment specifically designed for deep learning tasks. This system featured a state-of-the-art TESLA-V100 Graphics Processing Unit (GPU) of NVIDIA, boasting 32GB of dedicated memory. To unlock the full potential of the GPU, the environment utilized CUDA 11.0, a software toolkit that facilitates communication and optimization between the GPU and deep learning frameworks. Additionally, cuDNN 8.0, a library specifically designed to accelerate deep learning workloads on NVIDIA GPUs, was employed. Finally, the training environment itself was built on Python 3.10, a widely used programming language that offers a robust foundation for deep learning development. This combination of cutting-edge hardware, optimized software tools, and a modern programming language ensured a highly performant and efficient platform for conducting deep learning experiments.

5.3. Training, Validation and Testing Data

Three separate sets of training data were carefully selected: an 80% training set that was used to train the model, a 10% validation set that was used to track training progress and adjust hyperparameters, and a final 10% test set that was used to assess the model's generalizability on untested data objectively. This ensures the model learns effectively from the training data while avoiding overfitting, ultimately leading to robust performance on real-world applications. This 80-10-10 split is a common practice in machine learning, ensuring the model has sufficient data for training while also allowing for proper validation and testing to achieve robust performance.

- Training Set (80%): This will comprise 894 images (0.8 * 1118). This is the largest portion where the model learns by processing the data and identifying patterns.
- Validation Set (10%): This set will contain 112 images (0.1 * 1118). It is used to fine-tune hyperparameters and monitor the model's learning progress during training to prevent overfitting.
- Test Set (10%): This final collection, which will have 112 photos (0.1 * 1118), is meant to be used for an objective assessment of the model's performance using untested data. This aids in evaluating how well the model applies to actual situations.

5.4. Performance Evaluation Criteria

A set of standard metrics frequently used in object detection tasks to ensure an accurate and impartial evaluation of the suggested model's performance are taken into consideration in this research. F1-score, mean Average Precision (mAP), recall, precision, and precision are some of these. The percentage of expected positive detections that are actually present (proper targets) is precisely measured by precision [23][24]. It basically shows how well the favorable predictions made by our model worked. The precision metric measures the percentage of anticipated positive detections that are actually true positives. It reflects the accuracy of your model's positive predictions. A high precision value indicates that most of the objects of the model identified are indeed present in the image.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}\ X\ 100\% \quad (1)$$

$$\text{Average Precision (AP)} = \int_1^0 P(r)\ dr \quad (2)$$

Recall, also known as sensitivity, focuses on the completeness of the detections performed by the model [27]. It calculates the proportion of actual positive cases that your model correctly identifies. A high recall value signifies that the model captures most of the relevant objects in the image.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}\ X\ 100\% \quad (3)$$

The F1-score provides a harmonic mean to offer a balanced view of both metrics. It penalizes models that excel in either precision or recall at the expense of the other. A model that strikes a good balance between limiting false positives and identifying real positives is indicated by a high F1-score. Since a mAP takes into account the trade-off between precision and recall across various detection confidence thresholds, it is especially useful for object detection tasks [26]. It essentially calculates the average precision at various thresholds.

$$\text{Mean Average Precision (mAP)} = \frac{1}{n}\sum_{i=1}^{n} AP_I \quad (4)$$

5.5. Initialization Parameters of YOLO-NAS Network

Hyperparameters are variables whose values affect the learning process, and they define the model parameter values that a learning algorithm finally learns [27]. The accuracy of the model is also affected by the hyperparameter choices made to help an object identification model achieve its maximum accuracy [28]. Table 2 shows the parameters set for training the YOLO-NAS for crop and weed detection during the experiment. PPYoloELoss is used as a loss function [29].

5.6. Implementation Details and Challenges

Model training was computationally intensive, especially with large datasets. Additionally, dealing with imbalanced datasets and adapting the model to field conditions were significant hurdles which were addressed through custom data augmentation and model tuning techniques.

6. Results and Discussion

Loss functions play a critical role in guiding the training process of object detection models [30]. They measure the discrepancy between the predictions given by the model and the actual ground truth (real-world object locations and classifications). By minimizing these losses, the model learns to:

- Accurately classify objects: The classification loss penalizes the model for incorrectly assigning class labels to detected objects [31]. A lower classification loss indicates that the model can effectively distinguish between different object categories.
- Precisely localize objects: The localization loss measures how far the actual locations of objects deviate from the bounding boxes that are anticipated [32].

A reduced localization loss indicates that the model is good at enclosing the objects in the pictures with bounding boxes. Table 3 summarizes the classification loss, localization loss, and total loss for various object detection models considered for the experiment. SSD has the highest classification loss and a relatively higher localization loss compared to others. Faster RCNN performs better than SSD with MobileNetV2 in both classification and localization, indicating a more accurate detection overall. Faster RCNN is known for its accuracy but tends to be slower than other models like SSD and YOLO [33].

RetinaNet shows a good balance with a relatively low classification and the lowest localization loss among the non-YOLO models, indicating strong performance, particularly in pinpointing object locations. YOLOv8 significantly outperforms the previous models in both classification and localization loss [34]. YOLO (You Only Look Once) models are known for their speed and accuracy [35], which is reflected here. The proposed YOLO-NAS model shows the best performance among all listed models with the lowest classification and localization loss.

This indicates that this model is the most accurate in detecting and locating objects, making it highly efficient. Figure 4 shows the graph of total loss derived from YOLO-NAS during training. Our research investigates the critical task of crop and weed detection using deep learning techniques. We compared the performance of six distinct object detection models to identify the most effective approach for this agricultural application. These models were rigorously trained, validated, and tested on a comprehensive dataset of agricultural images.

Table 2 summarizes the specific models evaluated, along with their corresponding results on four key metrics: precision, recall, mean Average Precision (mAP), likely focusing on an Intersection over Union (IoU) threshold of 0.5 (mAP@0.5) (details in Table 4). By analyzing these metrics, we aimed to determine which model achieved the optimal balance between accurately identifying crops and minimizing the misdetection of weeds. SSD has good precision and a relatively high mAP, indicating a strong ability to identify objects correctly. The faster RCNN model has the lowest precision, recall, and mAP among the listed models. RetinaNet shows slightly higher precision than SSD with MobileNetV2 and similar recall.

Table 2. Parameters set for training the YOLO-NAS for crop and weed detection

Size of input images	Batch	Optimizer	Initial learning rate	Loss	Training steps
640 × 640	8	Adam	0.0001	PPYoloELoss	20,000

Table 3. The classification loss, localization loss, and total loss for various object detection models

Object Detection Models	Classification Loss	Localization Loss	Total Loss
SSD with MobileNetV2 (Quantized)	1.7971	0.2739	2.0710
Faster RCNN ResNet50 V1	0.9860	0.1327	1.1187
RetinaNet	0.9910	0.0787	1.0697
YOLOv8	0.2477	0.0578	0.3055
YOLO-NAS (Proposed)	0.2311	0.0427	0.2738

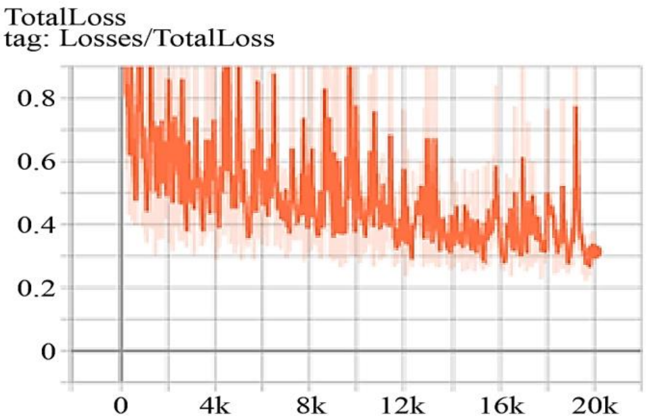

Fig. 4 Total loss obtained during training YOLO-NAS

Table 4. Performance Comparison of the proposed model with other state-of-the-art object detection model

Object Detection Model	Precision	Recall	mAP@0.50
SSD with MobileNetV2 (Quantized)	78.2	51.9	82.11
Faster RCNN ResNet50 V1	71.4	49.2	72.23
RetinaNet with ResNet50 V1	78.5	49.9	81.19
YOLOv8	81.2	53.3	85.56
YOLO-NAS (Proposed)	82.6	60.7	86.11

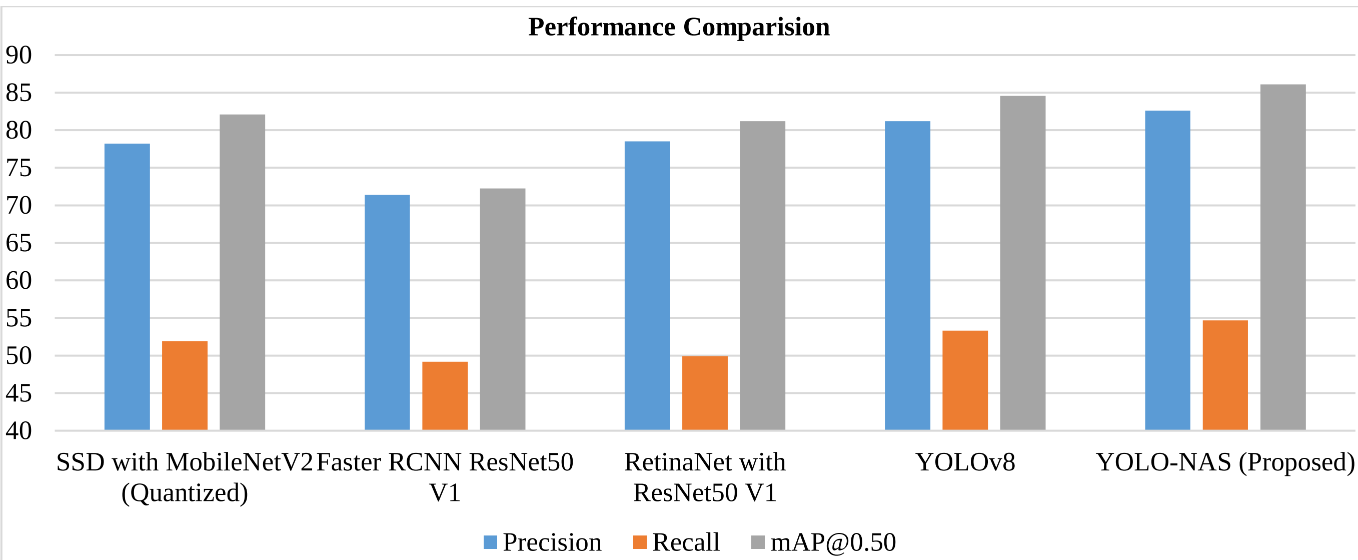

Fig. 5 Precision, Recall, and mAP@50 values show that YOLO-NAS outperforms all other state-of-the-art object detection models

Its mAP is also high, indicating robust detection performance, particularly in accurately identifying objects. YOLOv8 demonstrates high precision and recall, along with a high map. This suggests it is effective at both identifying and correctly locating objects, resulting in fewer false positives and false negatives. The proposed YOLO-NAS model outperforms all others in precision, recall, and mAP, as shown in Figure 5. This indicates that it is the most accurate and reliable model for detecting objects with minimal errors. In order to measure and analyse the accuracy of several object detection models, Intersection over Union (IoU) serves as a crucial metric [36]. Figure 7 showcases weed and crop objects detected by the proposed model. IoU helps quantify how well the model's predicted bounding boxes (boxes drawn around the objects) overlap with the actual locations of weeds and crops in the image. A higher IoU score indicates a greater overlap between the predicted and actual bounding boxes, signifying a more accurate detection of weeds and crops by the model. This information is valuable for tasks like targeted weed removal or automated crop yield estimation, where precise identification and localization of these objects are essential. Figures 6 (a) and (b) present the Real-time Weed and Crop Object Detection with IoU Visualization.

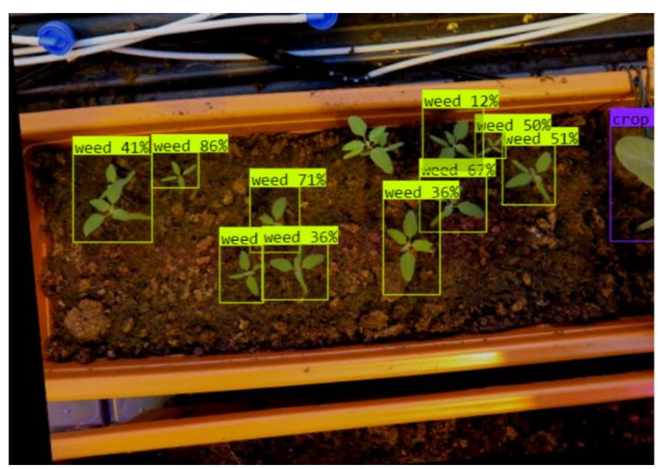

Fig. 6 (a) Real-time weed and crop object detection with IoU visualization – Sample 1

Fig. 6 (b) Real-time weed and crop object detection with IoU visualization – Sample 2. Figure (a) and (b) show images with bounding boxes for weed and crop classes. This visualization is crucial for informed weed removal and crop yield estimation.

Table 5. Summary of the major research carried out for crop and weed detection

Reference	Object Detection Model Used	mAP Score
I. Gallo (2023) [37]	YOLOv7	61.0%
Ahmad et al. (2021) [38]	YOLOv3	54.3 %
Gao et al. (2020) [39]	YOLOv3-tiny	82.9%
A. Rahman et al. (2023) [40]	Yolov5m	78.64%
Proposed Model	YOLO-NAS	86.11%

Table 5 shows the summary of the major research carried out for crop and weed detection. It shows that the proposed model achieved the highest mAP score of 82.9%, which outperforms all other models. In this research, the proposed model uses an automated search strategy to design the architecture, optimizing it for more efficient in detecting objects like weeds and crops. Moreover, YOLO-NAS balances the trade-off between speed and accuracy better than previous YOLO versions or other object detection models. The feature extraction techniques used in the proposed research capture finer details at multiple scales. The proposed model incorporates YOLO-NAS, which is specially designed for edge devices that, make it suitable for agricultural use cases where the model may need to run on drones, mobile devices, or other low-power systems directly in the field.

7. Future Directions and Improvements

The YOLO-NAS model for crop and weed identification may be improved, and this part examines new technologies that may help the model function better. Advances in NAS approaches, including more effective search algorithms or the integration of reinforcement learning for dynamic design adaption. Experimenting with new deep learning architectures such as transformers or hybrid CNN-Transformer models may also improve the robustness and accuracy of the model. Adding other data sources, such as multispectral or hyperspectral imaging, could greatly improve the model's capacity to distinguish between weeds and crops, particularly in difficult environmental circumstances. These developments could enhance the effectiveness of precision agriculture applications.

8. Conclusion

This research presents YOLO-NAS, a cutting-edge object detection model designed specifically for crop and weed detection in agricultural fields. By leveraging Neural Architecture Search (NAS) and advanced training techniques, YOLO-NAS achieves superior performance compared to existing state-of-the-art models. The experimental results demonstrate the ability of YOLO-NAS to identify and localize both crops and weeds accurately, outperforming competitors in terms of precision, recall, and mean Average Precision (mAP). With an mAP of 86.11%, YOLO-NAS maintains an optimal balance between correctly identifying crops and minimizing weed misdetection, addressing a critical challenge in precision agriculture. The proposed model offers a quantization-friendly design and deployment flexibility on various computational resources, from powerful workstations to resource-constrained edge devices, making it a versatile solution adaptable to diverse agricultural operations. Overall, YOLO-NAS represents a significant advancement in the field of automated crop and weed detection, offering a robust and efficient solution to enhance precision agriculture practices. Its seamless integration and deployment potential position it as a valuable tool for farmers, researchers, and agricultural technology companies, ultimately contributing to a more sustainable and productive agricultural sector.

References

[1] V.M. Abdul Hakkim et al., "Precision Farming: The Future of Indian Agriculture," *Journal of Applied Biology & Biotechnology*, vol. 4, no. 6, pp. 68-72, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[2] António Monteiro, Sérgio Santos, and Pedro Gonçalves, "Precision Agriculture for Crop and Livestock Farming—Brief Review," *Animals*, vol. 11, no. 8, pp. 1-18, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[3] Utilizing the Power of Data Analytics in Agriculture, Data Analytics, International Association Business of Analytics Certification, 2024. [Online]. Available: https://iabac.org/blog/utilizing-the-power-of-data-analytics-in-agriculture

[4] Bulbul Bamne et al., "Transfer Learning-Based Object Detection by Using Convolutional Neural Networks," *2020 International Conference on Electronics and Sustainable Communication Systems*, Coimbatore, India, pp. 328-332, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Thomas A Dorfer, YOLO-NAS: How to Achieve the Best Performance on Object Detection Tasks, Medium, 2023. [Online]. Available: https://towardsdatascience.com/yolo-nas-how-to-achieve-the-best-performance-on-object-detection-tasks-6b95347908d4
[6] Ye Mu et al., "A Faster R-CNN-Based Model for the Identification of Weed Seedling," *Agronomy*, vol. 12, no. 11, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[7] Nafeesa Yousuf Murad et al., "Weed Detection Using Deep Learning: A Systematic Literature Review," *Sensors*, vol. 23, no. 7, pp. 1-45, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[8] Muhammad Hammad Saleem, Johan Potgieter, and Khalid Mahmood Arif, "Weed Detection by Faster RCNN Model: An Enhanced Anchor Box Approach," *Agronomy*, vol. 12, no. 7, pp. 1-22, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[9] Vi Nguyen Thanh Le, Giang Truong, and Kamal Alameh, "Detecting Weeds from Crops under Complex Field Environments Based on Faster RCNN," *2020 IEEE Eighth International Conference on Communications and Electronics*, Phu Quoc Island, Vietnam, pp. 350-355, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[10] Hongxing Peng et al., "Weed Detection in Paddy Field Using an Improved RetinaNet Network," *Computers and Electronics in Agriculture*, vol. 199, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[11] J.M. López Correa et al., *8. Multi Species Weed Detection with Retinanet One-Step Network in a Maize Field*, Precision Agriculture '21, pp. 79-86, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[12] Sudheer Kumar Nagothu et al., "Weed Detection in Agriculture Crop Using Unmanned Aerial Vehicle and Machine Learning," *Materials Today: Proceedings*, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[13] Oluibukun Gbenga Ajayi, John Ashi, and Blessed Guda, "Performance Evaluation of YOLO v5 Model for Automatic Crop and Weed Classification on UAV Images," *Smart Agricultural Technology*, vol. 5, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[14] Jiqing Chen et al., "Weed Detection in Sesame Fields Using a YOLO Model with an Enhanced Attention Mechanism and Feature Fusion," *Computers and Electronics in Agriculture*, vol. 202, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[15] Velpula Sekhara Babu, and Nidumolu Venkatram, "Weed Detection and Localization in Soybean Crops Using YOLOv4 Deep Learning Model," *Traitement du Signal*, vol. 41, no. 2, pp. 1019-1025, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[16] Rohini Vaidya, Deci's YOLO-NAS: Next-Generation Model for Object Detection, 2023. [Online]. Available: https://medium.com/codex/decis-yolo-nas-next-generation-model-for-object-detection-8ccb7f2013a7
[17] Yolo-Nas, Ultralytics, 2023. [Online]. Available: https://docs.ultralytics.com/models/yolo-nas/
[18] Xiangxiang Chu, Liang Li, and Bo Zhang, "Make RepVGG Greater Again: A Quantization-Aware Approach," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 10, pp. 11624-11632, 2024. [CrossRef] [Google Scholar] [Publisher Link]
[19] Juan Terven et al., "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," *Machine Learning and Knowledge Extraction*, vol. 5, no. 4, pp. 1680-1716, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[20] Shuai Shao et al., "Objects365: A Large-Scale, High-Quality Dataset for Object Detection," *2019 IEEE/CVF International Conference on Computer Vision*, Seoul, Korea (South), pp. 8429-8438, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[21] YOLO for Object Detection, GUVI. [Online]. Available: https://forum.guvi.in/posts/6991/yolo-for-object-detection
[22] Kaspars Sudars et al., "Dataset of Annotated Food Crops and Weed Images for Robotic Computer Vision Control," *Data in Brief*, vol. 31, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]
[23] Touqeer Abbas et al., "Deep Neural Networks for Automatic Flower Species Localization and Recognition," *Computational Intelligence and Neuroscience*, vol. 2022, no. 1, pp. 1-9, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[24] Neeraj Kumar et al., "Leafsnap: A Computer Vision System for Automatic Plant Species Identification," *12th European Conference on Computer Vision – ECCV 2012*, Florence, Italy, pp. 502-516, 2012. [CrossRef] [Google Scholar] [Publisher Link]
[25] Alexis Joly et al., "Interactive Plant Identification Based on Social Image Data," *Ecological Informatics*, vol. 23, pp. 22-34, 2014. [CrossRef] [Google Scholar] [Publisher Link]
[26] Deval Shah, Mean Average Precision (mAP) Explained: Everything You Need to Know, V7labs.com, 2022. [Online]. Available: https://www.v7labs.com/blog/mean-average-precision
[27] Pasi Fränti, and Radu Mariescu-Istodor, "Soft Precision and Recall," *Pattern Recognition Letters*, vol. 167, pp. 115-121, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[28] Tanay Agrawal, *Hyperparameter Optimization in Machine Learning: Make Your Machine Learning and Deep Learning Models More Efficient*, Apress Berkeley, CA, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[29] Priyanka Kumari, Parking Space Detection Model Using PyTorch and Super Gradients, Labellerr, 2023. [Online]. Available: https://www.labellerr.com/blog/building-parking-space-detection-system-pytorch-super-gradients/
[30] Shang Jiang et al., "Optimized Loss Functions for Object Detection and Application on Nighttime Vehicle Detection," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 236, no. 7, pp. 1568-1578, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[31] Nhat-Duy Nguyen et al., "An Evaluation of Deep Learning Methods for Small Object Detection," *Journal of Electrical and Computer Engineering*, vol. 2020, no. 1, pp. 1-18, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[32] Rafael Padilla et al., "A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit," *Electronics*, vol. 10, no. 3, pp. 1-28, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Tiagrajah V. Janahiraman, and Mohamed Shahrul Mohamed Subuhan, "Traffic Light Detection Using Tensorflow Object Detection Framework," *2019 IEEE 9th International Conference on System Engineering and Technology*, Shah Alam, Malaysia, pp. 108-113, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[34] Mukaram Safaldin, Nizar Zaghden, and Mahmoud Mejdoub, "An Improved YOLOv8 to Detect Moving Objects," *IEEE Access*, vol. 12, pp. 59782-59806, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[35] Tausif Diwan, G. Anirudh, and Jitendra V. Tembhurne, "Object Detection Using YOLO: Challenges, Architectural Successors, Datasets and Applications," *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243-9275, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[36] Rafael Padilla, Sergio L. Netto, and Eduardo A.B. da Silva, "A Survey on Performance Metrics for Object-Detection Algorithms," *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, Niteroi, Brazil, pp. 237-242, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[37] Ignazio Gallo et al., "Deep Object Detection of Crop Weeds: Performance of YOLOv7 on a Real Case Dataset from UAV Images," *Remote Sensing*, vol. 15, no. 2, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[38] Aanis Ahmad et al., "Performance of Deep Learning Models for Classifying and Detecting Common Weeds in Corn and Soybean Production Systems," *Computers and Electronics in Agriculture*, vol. 184, pp. 1-30, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] Junfeng Gao et al., "Deep Convolutional Neural Networks for Image-Based Convolvulus Sepium Detection in Sugar Beet Fields," *Plant Methods*, vol. 16, pp. 1-12, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[40] Abdur Rahman, Yuzhen Lu, and Haifeng Wang, "Performance Evaluation of Deep Learning Object Detectors for Weed Detection for Cotton," *Smart Agricultural Technology*, vol. 3, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]