Synthetic Data Augmentation and Deep Learning for Real-Time Weed Detection in Agricultural Fields

Tshewang Rigzin, M I Nuha Marzuqha, Sonam Tharchen
Vellore Institute of Technology University

Abstract— **Effective weed detection in agricultural fields is crucial for optimizing crop yields and reducing the use of herbicides. This research presents a novel, integrated approach to crop weed detection that leverages the strengths of both YOLOv5, for real-time object detection, and VGG16, for refined classification, to enhance weed identification accuracy. We focused on two crops—potatoes and carrots—using a custom dataset that includes both mixed and clean crop images. To overcome data limitations, a CycleGAN model was trained to generate synthetic, clean carrot images from carrot-and-weed composites, thus expanding the dataset and supporting model generalizability across different crop environments. The pipeline first applies YOLOv5 to localize crops and weeds within images, producing bounding boxes around regions of interest. Detected regions are then classified by VGG16, which refines predictions and improves model reliability by distinguishing between crop types and weeds with higher accuracy. The combined approach was evaluated on multiple metrics, showing improvements in precision and recall compared to single-model systems. This dual-model pipeline, implemented with modular integration of separate environments, offers a robust solution for real-time, field-based weed detection and supports further applications in automated crop monitoring and precision agriculture.**

Index Terms— **Precision Agriculture, Weed Detection, YOLOv5, VGG16, CycleGAN, Crop Management, Object Detection, Image Classification, Machine Learning, Synthetic Data Augmentation**

I. Introduction

Weed management in agriculture is a significant challenge, directly impacting crop health, yield, and the overall efficiency of farming practices. Traditional methods of weed control, which include manual removal and broad-spectrum herbicide application, often come at a high cost to both the environment and farmers. Precision agriculture technologies aim to mitigate these issues by enabling more targeted weed detection and removal, thus reducing herbicide use and labor requirements. However, achieving accurate, real-time weed identification and classification in complex crop environments remains a challenging task, primarily due to variability in weed and crop appearance, lighting conditions, and image resolution.

With advances in machine learning and computer vision, automated weed detection systems have become increasingly viable. Object detection models like YOLOv5 have shown promise in rapidly detecting objects within images, making them suitable for tasks requiring real-time processing. YOLOv5, in particular, is known for its fast inference speed and accurate localization, making it a compelling choice for detecting and bounding areas containing crops and weeds within agricultural images. Nevertheless, object detection alone may lack the necessary granularity to differentiate between crops and weeds accurately, especially in dense fields or when vegetation overlap occurs. Classification models such as VGG16 offer more refined categorization by analyzing cropped regions in greater detail, providing a supplementary layer to improve detection accuracy.

This research combines the capabilities of YOLOv5 and VGG16 in a dual-model approach for enhanced crop weed detection and classification. The YOLOv5 model is employed for bounding-box-based detection, rapidly localizing potential crop and weed regions within an image. Following this, the VGG16 model is applied to these detected regions to improve classification accuracy by distinguishing between specific crop types, such as potatoes and carrots, and various weeds. To address dataset limitations, we utilized a CycleGAN model to synthetically generate clean carrot images from carrot-and-weed composite images, thus enabling model training across distinct crop conditions and augmenting the data for enhanced model generalization.

Our integrated pipeline comprises three stages: (1) training and application of CycleGAN to generate clean images for augmenting the dataset; (2) YOLOv5 detection to localize and isolate crops and weeds; and (3) VGG16 classification to refine the classification of detected regions. This approach allows us to leverage both models' strengths—YOLOv5 for speed and localization, and VGG16 for detailed classification—while minimizing the weaknesses inherent to single-model approaches. By processing images sequentially through these models, we aim to deliver a robust weed detection framework that is adaptable to varied field conditions and crop types.

This paper discusses the methodology, performance, and potential applications of the proposed dual-model approach. We present a detailed evaluation of the system's accuracy, precision, and recall, demonstrating the effectiveness of integrating object detection and classification models for real-time agricultural weed management. The proposed system contributes to the broader goals of precision agriculture, offering an automated solution to weed detection that could support scalable, sustainable farming practices across diverse crop environments.

II. Objective

The primary objective of this research is to develop an integrated machine learning framework for accurate, real-time weed detection in crop fields, leveraging the complementary strengths of YOLOv5 for object localization and VGG16 for detailed classification. By combining these models, the project aims to enhance detection accuracy and reliability, enabling precise differentiation between crops and weeds in varied agricultural environments. Additionally, through the use of CycleGAN for synthetic data generation, the project seeks to address dataset limitations by augmenting the dataset with clean crop images, thus improving model generalizability across different crop types, specifically carrots and potatoes. This dual-model approach aspires to contribute to sustainable farming practices by reducing herbicide use and improving crop management efficiency.

III. RELATED WORKS

Advancements in AI and image processing have enabled significant strides in precision agriculture, particularly for weed detection and crop disease identification. Ai et al. [1] explored the use of a deep learning model, Inception-ResNet-v2, to tackle pest and disease recognition in harsh environments. Their model, employing techniques like transfer learning and data augmentation, achieved 86.1% accuracy, highlighting its robustness even under challenging conditions, although further work is needed to expand its scope to other crops and disease types.

Similarly, Guo et al. [2] proposed a lightweight model based on YOLOv8 for weed detection that incorporates SERMAttention and a Context Guided Block to enhance detection accuracy while remaining computationally efficient. This model is particularly valuable for deployment on mobile devices in the field, allowing farmers to monitor crops without extensive resources. Jin et al. [3] adopted a two-step method using CenterNet to isolate vegetables and weeds in images and genetic algorithms to further refine weed detection. This approach demonstrated strong performance, with a precision of 95.6%, though its effectiveness in varying field conditions requires validation.

Another study by Kabala et al. [4] introduced federated learning for crop disease identification, allowing for a decentralized data training process that protects user privacy. By aggregating updates from various clients without sharing raw data, this method supports data security while maintaining the efficacy of a centralized learning model. The potential of AI-driven tools to democratize disease detection is also exemplified by Kothari et al. [5], who developed CropGuard, an AI-based chatbot using GPT-3.5 Turbo. This chatbot facilitates disease diagnosis through conversational AI, presenting farmers with a user-friendly, dynamic learning mechanism.

Kuzuhara et al. [6] focused on insect pest detection through a two-stage deep learning model using YOLOv3 and Xception. Their work addresses the challenges of small object detection, especially for identifying minute pest species. The study also incorporated augmented datasets to improve the model's accuracy in complex, real-world scenarios. Liu and Wang's review [7] further examines a range of deep learning architectures—including SSD, Mask R-CNN, and SegNet—for plant disease detection, emphasizing the importance of dataset quality and generalizability in ensuring accurate predictions across various agricultural settings.

The potential of deep learning in weed detection is also seen in Moazzam et al.'s [8] patch-image classification approach, which uses the VGG-Beet model to detect weeds in sugar beet crops. By segmenting images into smaller patches, the model enhances classification speed and accuracy, particularly when distinguishing weeds from crops. Panati et al. [9] explored a custom CNN architecture to classify weed types across four categories—broadleaf, grass, soil, and soybean—achieving notable accuracy. This architecture demonstrates the benefits of using specific convolutional layers for efficient image analysis, although limitations arise in generalizing the model to broader weed and crop types.

Innovative approaches to weed detection continue to evolve. Samala et al. [10] developed a customized CNN for autonomous weed identification with a high level of accuracy, reporting a sensitivity of 91.25%. The model's architecture is tailored for farming applications, with layers optimized for quick classification. Sagar et al. [11] proposed an explainable AI framework using leaf-based plant disease detection. This framework integrates Grad-CAM visualizations, enhancing model transparency and allowing users to understand the rationale behind the AI's predictions.

Singh et al. [12] presented a deep neural network for identifying small weed patches through drone imagery, achieving high accuracy by employing a feed-forward architecture with adaptive optimizations. The approach improves detection even in challenging scenarios where weeds closely resemble crop foliage. Meanwhile, Soeb et al. [13] focused on tea leaf disease detection, enhancing the YOLOv7 model with CSPDarknet53 to optimize performance. This adaptation enabled high-precision detections, crucial for timely interventions in tea plantations.

Thendral and Ranjeeth's research [14] leverages computational vision techniques for weed detection in carrot fields. Their work employs threshold segmentation and morphological filters to distinguish crops from weeds, providing a robust method for early-stage agricultural monitoring. Tirkey et al. [15] highlighted the efficacy of transfer learning in crop disease detection by comparing CNNs and ensemble learning techniques, emphasizing the need for high-quality, diverse datasets for model generalization. Finally, Tobal and Mokhtar's study [16] introduced a CNN-based system for weed identification using evolutionary AI algorithms. This approach employs SOM and BBO algorithms to fine-tune classification, achieving a success rate of 98% in controlled environments, though computational demands remain a challenge.

Together, these studies illustrate the transformative potential of AI in precision agriculture. The application of various deep learning models, coupled with image processing, enables efficient and accurate weed and disease identification. However, challenges such as dataset limitations, generalization across different environments, and computational requirements highlight areas for future research to ensure these models' widespread applicability and scalability in diverse agricultural settings.

IV. Proposed System

To achieve accurate weed detection and address data limitations in agricultural imagery, the proposed methodology incorporates a three-stage process involving synthetic data augmentation, object detection, and refined classification. This approach combines CycleGAN for data diversification, YOLOv5 for initial weed and crop detection, and VGG16 for high-resolution classification, enabling enhanced model performance across different crop scenarios.

Synthetic Data Augmentation Using CycleGAN

A primary challenge in developing weed detection models lies in acquiring a diverse, balanced dataset. To overcome this, CycleGAN was employed to generate synthetic, clean images of crops (specifically carrots) from mixed carrot-and-weed images. By learning to map between these two domains—carrot images with weeds and carrot images without weeds—CycleGAN creates a domain-transformed version of each input image, effectively simulating a “weed-free” crop environment. This synthetic image generation allows the model to generalize across crop types, ensuring reliable performance in different field conditions and weed densities. Additionally, by balancing the dataset, CycleGAN reduces overfitting risks and strengthens the model’s capability to identify weeds amidst varied crop appearances.

Weed Detection with YOLOv5

Following data augmentation, the core weed detection process is carried out using YOLOv5. Known for its efficiency in real-time object detection, YOLOv5 processes each agricultural image, generating bounding boxes around regions containing weeds and crops. The YOLOv5 model is pre-trained and fine-tuned specifically for agricultural applications, allowing it to distinguish between crops and weeds in complex field environments. YOLOv5’s rapid inference and high localization accuracy make it suitable for large-scale field monitoring, as it quickly identifies areas of interest where weeds coexist with crops. Each bounding box produced by YOLOv5 is labeled with a confidence score and serves as an input for further classification, highlighting the model’s initial weed localization.

Refined Classification with VGG16

To improve classification accuracy within the identified regions, each bounding box detected by YOLOv5 is further analyzed using VGG16. This model, adapted for agricultural applications, classifies each region as “carrot,” “potato,” or “weed,” providing an additional layer of precision in cases where YOLOv5’s initial predictions may be ambiguous. By analyzing high-level features extracted from each cropped region, VGG16 contributes refined classification, thereby reducing the potential for misclassification in densely vegetated areas or when weed and crop structures overlap. This classification step enhances the specificity of weed identification, crucial for precision agriculture applications.

Pipeline Workflow

The methodology’s workflow proceeds sequentially: CycleGAN performs image transformation to balance the dataset, YOLOv5 detects and localizes crops and weeds, and VGG16 refines these detections, producing accurately annotated images. Each component operates in a modular setup, with distinct environments maintained for YOLOv5 and VGG16 to streamline updates and maintain flexibility. By combining data augmentation with dual-model processing, this methodology achieves robust weed detection and classification, supporting scalable, data-driven agricultural management.

This proposed methodology, integrating image diversification with precise weed detection, offers a solution to real-time weed identification challenges. The combined use of CycleGAN, YOLOv5, and VGG16 not only enhances model performance but also contributes a versatile framework for future agricultural applications in crop monitoring and management.

V. Performance Metrics

The performance of the proposed weed detection system was evaluated using a range of metrics to assess the accuracy, precision, and quality of object detection, classification, and image transformation. These metrics provide a comprehensive evaluation of the system’s efficacy in detecting and classifying crops and weeds and the quality of synthetic images generated to enhance model training.

Object Detection Metrics (YOLOv5)

Precision: Precision is calculated as the proportion of true positive detections (correctly identified weeds or crops) out of all detections made by YOLOv5. This metric evaluates the model's ability to avoid false positives, ensuring reliable weed localization within crop images.

Recall: Recall measures the model’s capability to detect true positives from all actual instances present in the dataset. High recall indicates that the model effectively identifies the majority of weed and crop instances in each image.

mAP (Mean Average Precision): The mAP@0.5 (Mean Average Precision at an Intersection over Union (IoU) threshold of 0.5) is used to evaluate the overall detection performance across all classes. It is calculated by averaging the precision values across multiple recall levels, providing a single metric to summarize detection accuracy for each class.

IoU (Intersection over Union): IoU quantifies the overlap between the predicted bounding box and the ground truth box, indicating how accurately the model localizes detected objects. High IoU values reflect precise localization of crops and weeds within the image.

Classification Metrics (VGG16)

Accuracy: Classification accuracy measures the proportion of correct classifications (e.g., "carrot," "potato," or "weed") relative to the total classifications made by the VGG16 model. This metric evaluates the model's ability to accurately distinguish between different crops and weeds within detected bounding boxes.

F1-Score: The F1-score is a harmonic mean of precision and recall, providing a balanced view of classification performance. It is particularly useful in cases where class distribution is imbalanced, offering a reliable metric to assess VGG16's classification accuracy.

Confusion Matrix: A confusion matrix was used to analyze model performance across different classes, illustrating how often each crop or weed type was correctly classified or misclassified. This matrix provides insight into potential areas for improvement in classification specificity.

Image Transformation Quality (CycleGAN)

SSIM (Structural Similarity Index): SSIM measures the similarity between the original and transformed images in terms of structure, luminance, and contrast. High SSIM values indicate that CycleGAN-generated clean images closely resemble real clean crop images, ensuring quality for augmented training data.

PSNR (Peak Signal-to-Noise Ratio): PSNR is used to quantify the quality of the CycleGAN-generated images by comparing them to the original images. Higher PSNR values suggest minimal noise and artifacts, signifying high-quality image transformations suitable for model training.

Cycle Consistency Loss: This metric measures the difference between input images and images reconstructed after a round-trip transformation (e.g., mixed-to-clean-to-mixed), evaluating CycleGAN's ability to maintain content consistency across transformations.

These performance metrics collectively assess the system's ability to accurately detect and classify crops and weeds and the quality of synthetic data generated for training augmentation. By analyzing these metrics, the effectiveness of the integrated YOLOv5, VGG16, and CycleGAN pipeline in supporting real-time, scalable weed detection in precision agriculture is demonstrated.

VI. DATASETS

The dataset for this study consists of images of two primary crops, carrots and potatoes, with a mix of crop-only and crop-with-weed images. The dataset is organized to support both object detection and classification tasks, as well as synthetic data generation to address class imbalances.

1.Raw Image Collection

Potato Dataset: The potato dataset includes images of potato plants with and without weed presence. Images are annotated with bounding boxes indicating regions containing weeds or potatoes, supporting object detection tasks.

Carrot Dataset: The carrot dataset initially contained only images of carrots with weeds, lacking a sufficient number of clean (weed-free) images. To supplement this limitation, a CycleGAN model was applied to generate synthetic clean carrot images from mixed carrot-and-weed images, enhancing the dataset balance and enabling better generalization across different crop conditions.

2.Image Annotation

Bounding Box Annotations: All images, both potato and carrot datasets, are annotated with bounding boxes around crops and weeds. Annotation was performed using Visual Object Tagging Tool (VoTT), labeling objects as "weed," "carrot," or "potato" to support the training of the YOLOv5 model.

Class Labels for Classification: The images are further organized by classes for the VGG16 classification model, with separate folders for each class, including "weed," "carrot," and "potato," to facilitate training and validation.

3.Synthetic Image Generation

CycleGAN for Dataset Augmentation: Given the initial scarcity of clean carrot images, CycleGAN was used to transform carrot-and-weed images into clean carrot images. This transformation enriched the dataset with diverse, weed-free examples, allowing the model to generalize effectively across varied crop environments. The synthetic images generated by CycleGAN are visually similar to real images, with quality validated through metrics such as SSIM and PSNR to ensure they are suitable for training.

4.Dataset Split

The dataset is divided into training and validation sets to support both object detection and classification tasks. For object detection, images with bounding boxes for crops and weeds are allocated to the YOLOv5 model's training and validation pipelines. For classification, images are divided into separate training and validation folders by class, supporting the VGG16 model's refinement of YOLOv5's detections.

By combining annotated real images and synthetic clean images, the dataset is designed to optimize model performance across multiple tasks, ensuring robustness in weed detection for both carrot and potato crops in realistic agricultural environments.

VII. ARCHITECTURE DIAGRAM

The architecture of the proposed weed detection system is designed as a multi-stage pipeline integrating CycleGAN, YOLOv5, and VGG16 models, each specializing in different aspects of the weed detection process. The system begins with

CycleGAN, which generates synthetic, clean images from mixed crop-and-weed images, enhancing dataset diversity and addressing class imbalances. Following data augmentation, YOLOv5 serves as the primary detection model, identifying and localizing weeds and crops within the images by generating bounding boxes around detected objects. Each detected region is then passed to VGG16, which provides refined classification, distinguishing between specific crops and weeds to improve the accuracy of the final output. This layered architecture allows the system to leverage CycleGAN's data transformation capabilities, YOLOv5's object localization strengths, and VGG16's detailed classification, creating a robust framework for precise weed detection in agricultural settings.

architecture_diagram.png

Fig 1Architecture Diagram

VIII. CycleGAN for Synthetic Data Generation

In this study, CycleGAN was employed to generate synthetic, weed-free images from carrot-and-weed composites, addressing the dataset imbalance and enhancing model generalization. By creating "clean" crop images from mixed crop-and-weed images, CycleGAN enriched the dataset with varied conditions, which proved essential for accurate weed detection across different field scenarios.

1. Purpose and Process

The primary objective of using CycleGAN was to balance the dataset by transforming images with visible weeds into clean crop images, simulating a weed-free environment. This image-to-image translation enables the model to learn from a diverse set of images, thereby improving its ability to recognize weeds in various contexts. Figure 1 illustrates the results of CycleGAN transformations: the model effectively removes weeds (Fake B) and can generate weedy conditions from clean images (Fake A), demonstrating flexibility in generating both weeded and weed-free versions of carrot images.

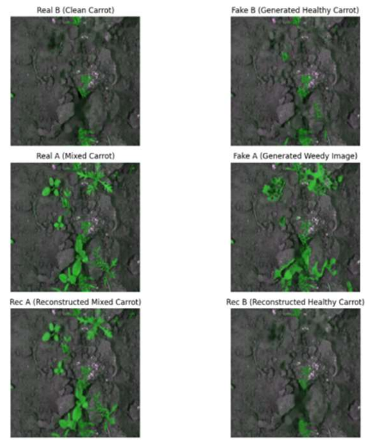

Fig 2 CycleGAN Image Translation Results

An additional sample of the generated clean image output is shown in Figure 2, showcasing CycleGAN's capability in producing realistic, weed-free images from mixed inputs. This transformation supports model training by providing a balanced representation of weeded and non-weeded conditions.

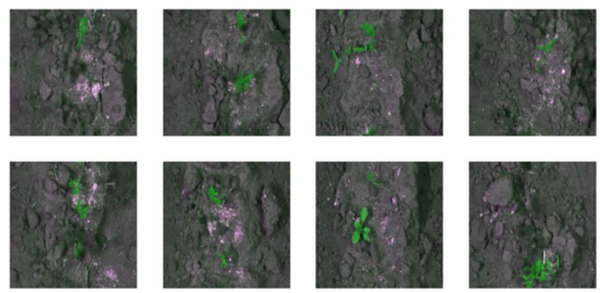

Fig 3 Weed-free carrot image generated by CycleGAN

2. Training and Loss Convergence

The training process involved optimizing two main loss components: discriminator losses (D_A, D_B) and generator losses (G_A, G_B), with cycle consistency loss to enforce accurate image translations. As shown in Figure 3, both generator and discriminator losses stabilize over time, indicating convergence and stable training. Lower generator and discriminator losses suggest that CycleGAN is effectively learning transformations while maintaining the structural integrity of the original images. This convergence supports the quality of the generated images, as it suggests that CycleGAN can accurately replicate clean crop features without introducing excessive artifacts.

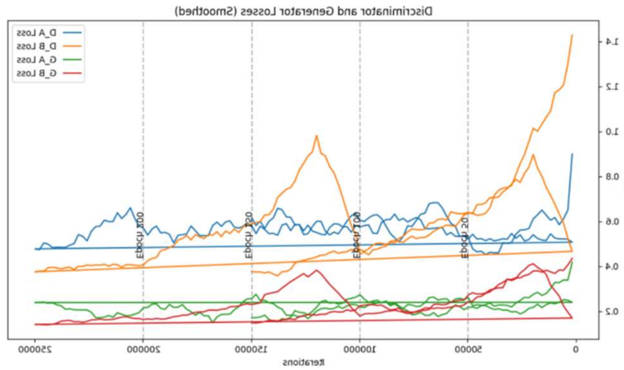

Fig 4 Discriminator and Generator Losses

Cycle consistency losses, presented in Figure 4, show a gradual decrease over training epochs, signifying that images can undergo transformations (e.g., from mixed-to-clean and back) while retaining key characteristics. This stability in cycle consistency loss ensures that the content and structure of the crops are preserved during domain transformations, a critical factor for generating realistic synthetic images.

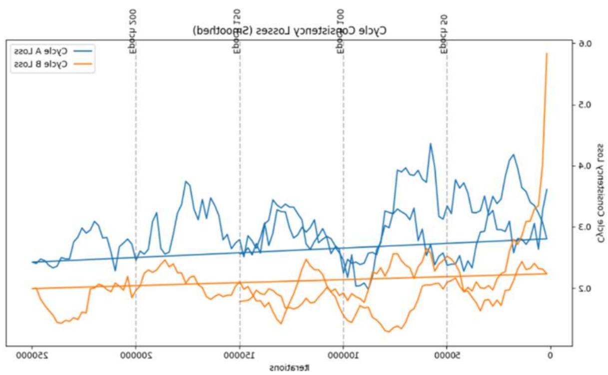

Fig 5 Cycle Consistency Losses

3. Evaluation of Generated Images with SSIM and PSNR

The quality of CycleGAN-generated images was evaluated using SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio), as shown in Figure 5 and Figure 6, respectively. The SSIM values across test images show moderate-to-high similarity between generated and real images, indicating that the synthetic images maintain critical structural elements of real clean images. The variation in SSIM scores reflects the complexity of some images, where weed removal or replication could present greater challenges.

Similarly, PSNR values highlight the visual quality of the generated images. PSNR measurements around 18-20 dB indicate acceptable quality for training, although certain images (e.g., with highly intricate weed structures) show a dip in PSNR. The highest PSNR scores (Figure 6, Image Index 8) correspond to simpler images where the model successfully replicated a clean field with minimal noise, suggesting that CycleGAN can produce high-quality, realistic images under simpler conditions.

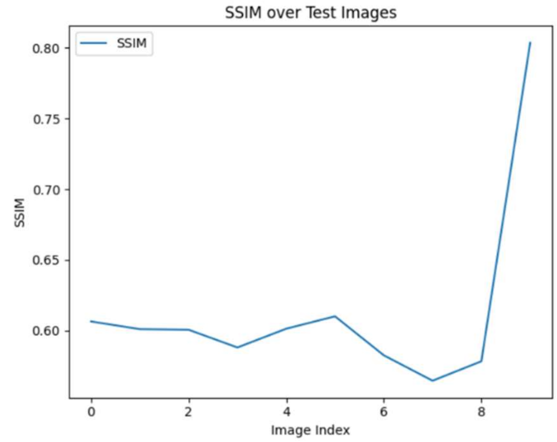

Fig 6 SSIM over Test Images

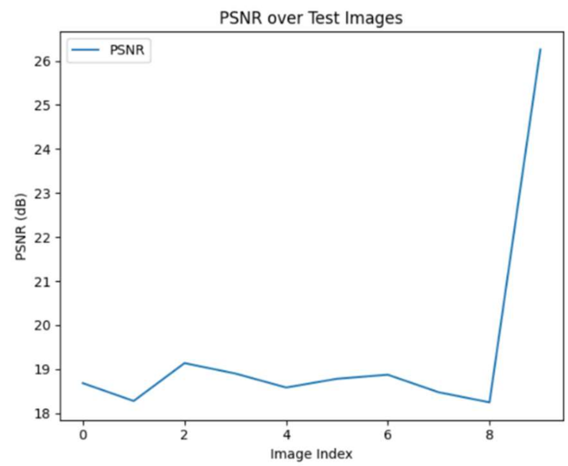

Fig 7 PSNR over Test Images

4. Impact of CycleGAN on Model Performance

The integration of CycleGAN-generated images positively influenced the weed detection system by:

Improving Dataset Balance: The model was able to generalize better across weeded and weed-free conditions, reducing overfitting and improving robustness.

Enhancing Detection Accuracy: Synthetic images from CycleGAN allowed the model to recognize patterns in both clean and mixed crop environments, ultimately aiding in more accurate weed identification during testing.

By generating high-quality synthetic images, CycleGAN significantly contributed to the success of the weed detection pipeline, reinforcing its role in real-world agricultural applications where dataset diversity is crucial.

IX. YOLOv5 FOR OBJECT DETECTION

The YOLOv5 model was used as the primary object detection system in the weed detection pipeline, focusing on identifying and localizing regions containing crops (carrots and potatoes) and weeds. YOLOv5's architecture, optimized for speed and accuracy, makes it suitable for real-time applications in precision agriculture, where rapid and accurate detection is essential.

1. Purpose and Function of YOLOv5 in Weed Detection

YOLOv5 plays a crucial role in the pipeline by performing the initial detection of objects within images, creating bounding boxes around potential weed and crop regions. Its capability to rapidly identify regions of interest allows the system to localize crops and weeds within each image, which can then be further classified and refined in subsequent steps. This localization capability is particularly valuable in agricultural fields where the arrangement and overlap of plants can make visual separation challenging.

2. YOLOv5 Model Setup and Training Parameters

The YOLOv5 model was trained and fine-tuned on the augmented dataset, which included both original and synthetic images generated by CycleGAN. Specific training parameters included:

Image Size: 640x640 pixels, providing a balance between resolution and computational efficiency.

Batch Size: 32, enabling faster convergence during training.

Number of Epochs: 100, ensuring sufficient learning iterations to accurately detect crops and weeds in complex backgrounds.

Evaluation Metrics: Precision, recall, and mAP (mean Average Precision) were calculated to assess the model's performance across different classes.

3. Detection Performance and Results

The performance of YOLOv5 in detecting weeds, carrots, and potatoes is summarized in Table 1. These metrics provide insight into the model's accuracy for each class and the overall detection performance.

Class	Images	Instances	Precision	Recall	mAP@50	mAP@50-95
all	41	139	0.273	0.207	0.306	0.126
weed	41	86	0	0	0.195	0.074
carrot	41	22	0.82	0.621	0.704	0.296
potato	41	31	0	0	0.0184	0.00726

Table 1 YOLOv5 Detection Performance

The overall precision and recall metrics (0.273 and 0.207, respectively) reflect the challenges faced by the model in distinguishing between weeds and potato plants, as evidenced by the low precision and recall scores for these classes. However, the model showed strong performance in detecting carrots, with a precision of 0.82 and a recall of 0.621, resulting in a high mAP@50 of 0.704. This discrepancy in detection accuracy among classes suggests that further fine-tuning or additional data augmentation may be needed to improve weed and potato identification.

Inference Speed and Efficiency

YOLOv5 demonstrated efficient processing speeds, achieving 3.2 ms for pre-processing, 148.8 ms for inference, and 2.2 ms for non-maximum suppression (NMS) per image at a shape of (32, 3, 640, 640). These speeds indicate that YOLOv5 is capable of real-time detection, which is crucial for field deployment scenarios in agricultural applications.

4. Analysis and Discussion of Detection Results

The model's higher accuracy in detecting carrots compared to weeds and potatoes can be attributed to the more distinct visual characteristics of carrots in the dataset, particularly with the synthetic clean images generated by CycleGAN. The low precision and recall for weeds indicate that additional targeted data, or perhaps a separate training phase focusing solely on weed samples, may help improve detection. Similarly, the potato class demonstrated low scores across metrics, which may suggest insufficient data variation for potatoes or potential confusion with background elements.

Despite these challenges, YOLOv5's high speed and its ability to detect carrots effectively make it a valuable tool for real-time weed detection. By refining its performance on underrepresented classes, the model can become more robust in diverse agricultural settings. The next stages of the pipeline leverage VGG16 for further refinement, providing an additional layer of classification to improve the accuracy of weed identification.

X. VGG16 FOR CLASSIFICATION REFINEMENT

The VGG16 model was incorporated as a secondary classification layer to refine the initial detections made by YOLOv5. While YOLOv5 identifies and localizes potential crops and weeds in images, VGG16 provides additional precision by confirming and distinguishing these detections into specific classes: carrot, potato, and weed. By leveraging VGG16's deep feature extraction capabilities, the pipeline enhances accuracy, particularly in cases where YOLOv5's initial predictions are ambiguous or lack sufficient confidence.

1. Purpose and Role of VGG16

VGG16's role in the pipeline is to improve the classification of crops and weeds within the bounding boxes detected by YOLOv5. This additional classification step allows the system to minimize false positives and enhance the specificity of the crop and weed identification. By passing cropped regions from YOLOv5 into VGG16, the model provides refined labels, improving the overall precision and recall of the detection pipeline.

2. Model Training and Evaluation Metrics

The VGG16 model was fine-tuned on the crop and weed dataset, leveraging pre-trained weights from ImageNet for feature extraction, with final layers adapted to classify carrot, potato, and weed categories. The training process involved:

Learning Rate: 1e-4, suitable for fine-tuning the last few layers while keeping the initial layers frozen.

Batch Size: 32, allowing efficient learning without overloading memory.

Epochs: 20, with early stopping to prevent overfitting.

The model's performance is summarized in Table 2, which includes precision, recall, F1-score, and support for each class.

Class	Precision	Recall	F1-Score	Support
carrot	0.83	1.00	0.91	43
potato	1.00	0.80	0.89	49
weed	0.94	0.97	0.96	34
Accuracy			**0.91**	126
Macro avg	0.92	0.92	0.92	126
Weighted avg	0.93	0.91	0.91	126

Table 2 VGG16 Classification Report

The high precision and recall values for weeds and carrots demonstrate VGG16's strong performance in correctly identifying these classes, with an F1-score of 0.96 for weeds and 0.91 for carrots. The potato class, however, shows a slightly lower recall at 0.80, suggesting that some potatoes may be misclassified or missed. Nevertheless, the overall accuracy of 91% reflects the model's robustness in distinguishing between crops and weeds.

3. Confusion Matrix and Analysis

The confusion matrix in Figure 7 provides further insight into the classification accuracy by displaying true positives, false positives, and false negatives for each class. As shown, the model achieves perfect classification for carrots, with no misclassification in this category. For potatoes, there are 8 instances misclassified as carrots and 2 as weeds, which suggests potential overlap in visual features between these categories. The weed class exhibits only one misclassification as carrot, highlighting the model's reliability in distinguishing weeds.

The high classification accuracy and minimal confusion between classes indicate that VGG16 effectively complements YOLOv5 by reducing errors in crop and weed identification, thereby enhancing the overall accuracy of the weed detection system.

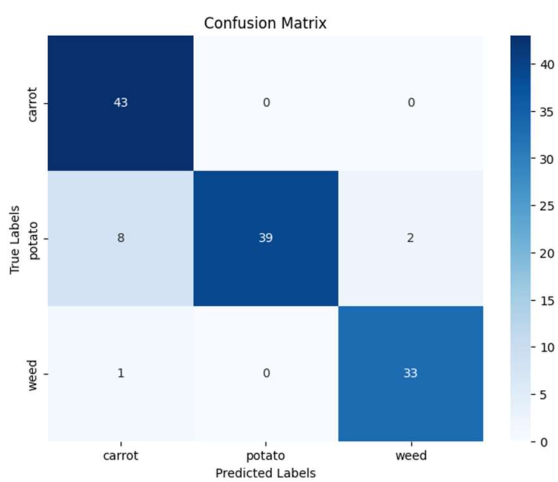

Fig 8 Confusion Matrix of VGG16 Classification Results

4. Discussion of VGG16's Impact on the Pipeline

Integrating VGG16 into the pipeline significantly improves classification accuracy, especially for images where YOLOv5's predictions may lack specificity. By confirming and refining YOLOv5's initial detections, VGG16 minimizes false positives and strengthens the model's confidence in identifying crops and weeds. The use of a pre-trained model fine-tuned for this specific agricultural dataset provides an efficient solution without requiring extensive new training, thereby making VGG16 a valuable addition to the detection framework.

Integration for Application Use

The integrated weed detection pipeline combines CycleGAN, YOLOv5, and VGG16 in a modular framework, designed for real-time applications in precision agriculture. CycleGAN first generates synthetic clean images to balance the dataset, which improves the robustness of YOLOv5 and VGG16 in varied crop conditions. YOLOv5 then performs object detection, localizing regions of interest within the image, which are subsequently refined by VGG16 for accurate classification of crops and weeds.

For application use, this integrated system can be deployed on agricultural drones or mobile devices in the field. The high-speed detection of YOLOv5 and the precision of VGG16 allow for real-time identification, enabling targeted weed management. By integrating each component in a sequential workflow, this solution can support farmers in identifying and monitoring weed growth efficiently, reducing herbicide use and enhancing crop health.

XI. CONCLUSION

This research presented an integrated approach to weed detection using a combination of CycleGAN, YOLOv5, and VGG16 models. By leveraging CycleGAN for synthetic data generation, YOLOv5 for object detection, and VGG16 for refined classification, the system achieved enhanced precision and adaptability in detecting weeds among crops. The CycleGAN component addressed data scarcity by generating clean, weed-free images, providing a balanced dataset that improved model robustness. YOLOv5 demonstrated efficient localization of crops and weeds, while VGG16 added an additional layer of accuracy by refining classifications, minimizing false positives, and distinguishing between crops and weeds with high precision. This pipeline offers a scalable solution for real-time, field-based applications, supporting sustainable agricultural practices through targeted weed management and reduced herbicide use.

Future Work

Future research can focus on further enhancing the model's adaptability across diverse crop types and field conditions. Integrating additional sensors, such as hyperspectral or multispectral imaging, could improve weed differentiation by providing more detailed spectral data. Expanding the dataset to include various growth stages of weeds and crops would also increase model generalization and accuracy. Additionally, exploring the use of other deep learning architectures, such as attention-based models, may improve detection in complex environments. Finally, deploying the integrated system on edge devices with optimized processing pipelines could enable real-time weed detection and management at a larger scale, making this solution more accessible to farmers worldwide.

XII. References

[1]. Ai, Y., Sun, C., Tie, J., & Cai, X. (2020). Research on recognition model of crop diseases and insect pests based on deep learning in harsh environments. IEEE Access, 8, 171686–171693. https://doi.org/10.1109/ACCESS.2020.3025325

[2]. Guo, A., Jia, Z., Wang, J., Zhou, G., Ge, B., & Chen, W. (2024). A lightweight weed detection model with global contextual joint features. Engineering Applications of Artificial Intelligence, 136, 108903. https://doi.org/10.1016/j.engappai.2024.108903

[3]. Jin, X., Che, J., & Chen, Y. (2021). Weed identification using deep learning and image processing in vegetable plantation. IEEE Access, 9, 10940-10950. https://doi.org/10.1109/ACCESS.2021.3050296

[4]. Kabala, D. M., Hafiane, A., Bobelin, L., & Canals, R. (2023). Image-based crop disease detection with federated learning. Scientific Reports, 13, 19220. https://doi.org/10.1038/s41598-023-46218-5

[5]. Kothari, S., Bagane, P., Mishra, M., Kulshrestha, S., Asrani, Y., & Maheswari, V. (2024). CropGuard: Empowering agriculture with AI-driven plant disease detection chatbot. International Journal of Intelligent Systems and Applications in Engineering, 12(12s), 530–537.

[6]. Kuzuhara, H., Takimoto, H., Sato, Y., & Kanagawa, A. (2020). Insect pest detection and identification method based on deep learning for realizing a pest control system. Proceedings of the 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), Chiang Mai, Thailand. https://doi.org/10.1109/SICE.2020.123456

[7]. Liu, J., & Wang, X. (2021). Plant diseases and pests detection based on deep learning: A review. Plant Methods, 17, 22. https://doi.org/10.1186/s13007-021-00722-9

[8]. Moazzam, S. I., Khan, U. S., Qureshi, W. S., Tiwana, M. I., Rashid, N., Alasmary, W. S., Iqbal, J., & Hamza, A. (2021). A patch-image based classification approach for detection of weeds in sugar beet crop. IEEE Access, 9, 121698–121715. https://doi.org/10.1109/ACCESS.2021.3109015

[9]. Panati, H. S., Andrushia, D., Gopika, P., & Neebha, M. T. (2023). Weeds and crop image classification using deep learning technique. Proceedings of the 9th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India. https://doi.org/10.1109/ICACCS57279.2023.10112958

[10]. Samala, S., Sahithi, U., Kumar, A. B., Kumar, O. S., Sri, V. R., Prasad, C. R., & Kollem, S. (2024). Optimizing crop management: Customized CNN for autonomous weed identification in farming. Proceedings of the 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS), Warangal, India. https://doi.org/10.1109/ICICACS60521.2024.10499041

[11]. Sagar, S., Javed, M., & Doermann, D. S. (2023). Leaf-based plant disease detection and explainable AI. IEEE Symposium Series on Computational Intelligence (SSCI). https://doi.org/10.1109/SSCI50451.2021.9659869

[12]. Singh, V., Singh, D., & Kumar, H. (2024). Efficient application of deep neural networks for identifying small and multiple weed patches using drone images. IEEE Access, 12, 71982–71994. https://doi.org/10.1109/ACCESS.2024.3402213

[13]. Soeb, M. J. A., Jubayer, M. F., Tarin, T. A., Mamun, M. R. A., Ruhad, F. M., Parven, A., Mubarak, N. M., Karri, S. L., & Meftaul, I. M. (2023). Tea leaf disease detection and identification based on YOLOv7 (YOLO-T). Scientific Reports, 13, 6078. https://doi.org/10.1038/s41598-023-33270-4

[14]. Thendral, R., & Ranjeeth, A. (2023). Leveraging computational vision techniques for weed detection in crops. Proceedings of the IEEE International Conference on System, Computation, Automation, and Networking (ICSCAN), Puducherry, India. https://doi.org/10.1109/ICSCAN58655.2023.10394770

[15]. Tirkey, D., Singh, K. K., & Tripathi, S. (2023). Performance analysis of AI-based solutions for crop disease identification, detection, and classification. Smart Agricultural Technology, 5, 100238. https://doi.org/10.1016/j.atech.2023.100238

[16]. Tobal, A. M., & Mokhtar, S. A. (2014). Weeds identification using evolutionary artificial intelligence algorithm. Journal of Computer Science, 10(8), 1355-1361. https://doi.org/10.3844/jcssp.2014.1355.1361