International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

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

M I Nuha Marzuqha[1], Tshewang Rigzin[2], Sonam Tharchen[3]

[1,2,3]Vellore Institute of Technology, Chennai

Abstract

The ability to identify weeds in agricultural fields is critical to increasing productivity of crops as well as minimizing the use of herbicides. This paper presents a combined approach focusing on crop-weed discrimination employing the object detection capabilities of YOLOv8, the image classification power of VGG16 and, Grad-CAM, for the purpose of understanding the concerns of the model predictions. With a custom dataset of images consisting of mixed and clean images of potato and carrot crops, enhanced data limitations are achieved by training a CycleGAN model to synthesize clean carrot images from carrot-weed composite images, thus augmenting the dataset and allowing for better model performance in different settings. The pipeline begins with a vision object detection network called YOLOv8, which is used to detect the crops and the weeds in the image by drawing bounding boxes around the areas of interest. VGG16 then takes it a step further by classifying the regions, specifically differentiating between crops and weeds even more accurately. Classification outcomes are further enhanced by Grad-CAM which helps to visualize and elucidate the classifications giving an understanding of an area of interest in the models' prediction. when assessed across the various metrics, the combined strategy improves the precision and recall measures over the single model systems. This two-model modular design avails a suitable approach for weed detection in the fields in a real-time situation and can be extended for crop monitoring and precision agriculture purposes.

Keywords: Precision Agriculture, Weed Detection, YOLOv8, VGG16, CycleGAN, Grad-CAM, Crop Management, Object Detection, Image Classification, Machine Learning, Synthetic Data Augmentation.

Introduction

Weed management in agriculture is a tough nut to crack because it affects the growth, health, and yield of crops, as well as the entire farming process. Discreet weeding techniques (like hand-pulling or using herbicides) tend to be too expensive. Processed agriculture technologies seek to solve this breed of problems by allowing the detection and removal of weeds in a more focused manner than herding limitations and herbicide usage. Only, proper discrimination and classification of real-time and in-situ weed images in objective and complex crop scenarios remains a challenge to address variance of weed's and crop's appearance, light conditions and image resolution.

Because of the recent innovations in machine learning and computer vision, the imperative to have an automated weed detection system is becoming more practical. Detectors such as YOLOv8 have been able to perform effectively in real time processing tasks such as high speed object detection. Since YOLOV8 is the architecture that boasts of better speed when inferring regions of interest and improved accuracy in localization of such regions, it is well poised to detect and draw bounding boxes for cases with crops and weeds in the images from the field. Object detection in this case is somewhat limiting in that it may not provide the level of detail needed to accurately discriminate between a weed and a crop i.e. the level of discrimination is not enough when the crops are very close or intermixed with the weeds. Classification models such as VGG16 provide enhanced classification after more rigorous scrutiny of previously cropped sections hence improved detection.

In this work, we combine the advantages of both YOLOv8 and VGG16 in a two-model system for detecting and classifying crops and weeds more effectively. The identification part involves bounding box detection using YOLOv8 in which crop and weed regions in an image are efficiently identified. Thereafter, these regions are processed using VGG16 aimed at improving the classification of different crops (potato and carrot) and weed species. Due to the lack of adequate datasets, we therefore use a CycleGAN model to create clean carrot images from composite images of carrot and weed thus augmenting the dataset for better model training generalization.

Our overall framework consists of four stages: (1) CycleGAN training and use for producing the synthetic clean images which improves the dataset; (2) Crop and weeds localization using YOLOv8 detection; (3) further classification of the localized regions using VGG16; and (4) Grad-CAM to explain which aspects of the input the model focused on while performing the classification, thereby assisting detection and classification. This systems architecture is designed to utilize YOLOv8 for Laplacian co-ordinate localization in real-time, VGG16 to enable hierarchy of classification while Grad-CAM also assists in interpretation of this expansion by outlining the features that contributed most to the given classification. This paper illustrates the strategy, the effectiveness and the possible uses of this two systems combined approach. We analyze performance in terms of accuracy, precision and recall of the system, thus showing the feasibility of real-time indoor weed management using object detection and classification systems. Such system is highly correlated with the ideas embedded in precision agriculture as it provides a means of weed management that is technologically driven and can be deployed across various crops and production systems.

Objective

The main aim of the research is to create a machine learning system for accurate and real-time weed identification within agricultural fields. This will take advantage of the object localizing capabilities of YOLOv8 and the image classification power of VGG16. The project intends to overcome the challenges of accurate target detection and target discrimination in complex rice field environments by fusing these models. In addition, the project aims to use CycleGAN for creating synthetic data in order to overcome the challenges in the dataset and enhance it with clean images of crops for model generalizability targeting crop types such as carrots and potatoes. The Grad-CAM allows for a better understanding of the architecture by visualizing the model's predictions, which increases trust in the detection system. Such a dual model works against the overuse of herbicides to support efficient crop management.

Literature Survey

The deployment of modern technologies such as Artificial Intelligence and imaging techniques has led to the development of new methods of farming, especially when it comes to recognizing and preventing weeds and plant diseases. The possibility of recognizing pests and diseases in extreme conditions through deep learning model Inception-ResNet-v2 has been researched by Ai et al. [1]. When employing transfer learning, data augmentation, and so on, the model reached 86.1% accuracy so demonstrating a robust performance during extreme circumstances, however, the model is not yet ready to be applied on other crops or types of diseases.

Guo et al. [2], on the other hand, have introduced a lightweight and efficient weed detection model based on YOLOv8 and enhanced by SERMAttention and a Context Guided Block aiming at improving detection accuracy. Such a model is critical in the course of the studies since it can be run on mobile devices in the field and thus enabling crop monitoring even with low computer resources. For instance, Jin et al. [3] used CenterNet at first, to separate images of vegetables from those of weeds and secondly employed genetic algorithms to enhance the process of weed identification. It performed quite well achieving 95.6% precision but the use of such technique has to be tested in different field conditions.

In another case, Kabala et al. [4] inserted the ability to use federated learning for identification of crop diseases which allows training even without collecting data from users. Focused on obviating the need of sharing actual data among the clients, this approach allows aggregation of client updates and still retains the advantages of centralized training. The possibilities offered by Kothari et al. [5] in the case of an AI tool enhancing disease detection are not limited to advanced algorithms as their team built a CropGuard chatbot based on GPT-3.5 Turbo. With this chatbot, it onyee datanosis solutions

Kuzuhara et al. [6] analyzed methods of detecting insect pests using a two-stage deep learning model, YOLOv3 and Xception. This work confronts challenges such as small object detection whereby minute pest species are targeted. The augmented datasets were part of the study to enhance model performance in real-life scenarios where pests are not usually found. It also provides an overview of Liu and Wang's review [7], which is another literature on the application of deep learning for plant diseases using architectures such as SSD, Mask R-CNN, SegNet, and many other similar ones, where the focus is on datasets, their quality, and how broad the scope of the dataset is to enable correct predictions in different farms.

A few works on the use of deep learning in weed detection have also been noted, especially in Moazzam et al. [8] who used the VGG-Beet model for patch-image classification of sugar beet crops. This method is effective in enhancing the classification of different crops and weeds, especially when the images are divided into patches. Panati et al. [9] used a custom CNN architecture to classify weeds in four categories: broadleaf, grass, soil, and soybean, with considerable accuracy. This aspect presents the advantages of specific convolutional layers during image processing tasks; however, this poses a challenge when applying the model to other types of crops and weeds that fall outside the studied range.

New methods on weed detection continue to be developed. Samala et al. [10] provided a customized CNN for automatic identification of weed species that turned out to be very efficient as they reported a sensitivity of 91.25%. The architecture of the model is more suited for agricultural purposes as it has layers that facilitate faster classification. Sagar et al. [11] introduced an explainable AI framework that used leaf based plant disease detection. This framework is supplemented with Grad-CAM visualizations and improvements as well.

Adoption of a feed-forward deep neural network with various optimizations Singh et al. [12] was able to use drone images to recognize and classify small patches of weed with a great amount of accuracy. The approach is also beneficial when detection proves to be quite difficult, especially within crop fields where weeds look like crops. Soeb et al. [12] on the other hand, presented a modified version of YOLOv7 with

CSPDarknet53 that was used to tea leaf disease classification. This proved to be effective as precision detection was very important for timely mitigation of threats to the tea crop.

The work of Thendral and Ranjeeth [14] focuses on weed detection from carrot vegetables using computer vision techniques. It uses thresholding and morphological filters to separate crops and weeds, therefore offering a solution for effective monitoring of agriculture especially at early stages of growth. In the study, Tirkey, et al., [15]. discussed how useful transfer learning is in the area of crop diseases containment as CNNs as well as pooled learning approaches were put to use, we also wish to mention how crucial good quality and variated datasets is in addressing the issue of deep learning models put under the scope. As for the study by Tobal and Mokhtar – a CNN system for weed control helped with built evolutionary AI – optimum weed identification system was presented. SOM and BBO algorithms assist in enhancing classification and attain a degree of success of 98% when the environments are controlled, although still it poses difficulty in terms of computation.

Altogether these works emphasized how impactful AI technology is to modern agriculture, particularly in the precision farming sector. The use of different deep learning models along with image processing facilitates quick and correct identification of weeds and diseases. Nevertheless, there are still issues like, the datasets are not sufficient, it is hard to generalize in other fields, there are requirements for computing power that call for more work in the future in order to permeate these models with a wider scope.

Proposed System

All these factors motivate the need for the integration of a four-stage approach that comprises synthetic data augmentation, object detection, elevated garbage classification, and model interpretation so that effective weed detection will be performed and the existing limitations of agricultural imagery data will be dealt with. The four components work together within this framework: synthetic data exploitation through CycleGAN, initial weed and crop detection using YOLOv8, high approval-rate classification with VGG16, and lastly, for model trust, Grad-CAM is applied to enhance the performance of the model in different crops scenarios.

Synthetic Data Augmentation Using CycleGAN

One of the major obstacles faced when designing a system for weed detection, is the collection of an adequate and well-balanced sample. In order to solve this problem, synthetic images of crops, namely carrots, with no weeds present were created by means of CycleGAN from pictures containing both carrots and weeds. CycleGAN works by generating an image from the target domain given an image in the source domain with the mapping being learnt when the two images are of different domains such as carrot images with weeds and images of only carrots without any weeds. In this manner, the input image is synthesized to look like the input image with a specific crop but without the extraneous crops (weeds). Such generation of a synthetic image helps the model to extend over the crops and not only depend on a single type of crop to perform effectively in the fields even when the crops have different conditions and different levels of weeds. On top of this, such a balancing and overfitting control helps in utilizing the dataset more effectively to avoid overfitting in the model and improves the performance of the model to help in distinguishing the weeds when they grow among other crops that appear completely different to the crops of interest in the centre.

Weed Detection with YOLOv8

Data augmentation is performed first and then the primary weed detection process is executed using YOLOV8. Given its capability in real-time object detection, YOLOv8 is used for agriculture image processing where bounding boxes are drawn around weed as well as crop images. The YOLOv8 model is pretrained and finetuned for agriculture purposes, thus making it efficient in differentiating crops mean and weed in complex field settings. Since YOLOv8 has fast inference and high localization accuracy, it is effective for large field monitoring as it assesses regions of interest containing crops and weeds in a matter of seconds. All the bounding boxes generated by YOLOv8 come accompanied with confidence scores which are utilized to classify the further, this marking the first stage of the model in weed localization.

Refined Classification with VGG16

To enhance accuracy in classification of the crops within the given regions, each bounding box produced by YOLOv8 is also processed by VGG16. Region after region, each section is designated 'carrot', 'potato' or 'weed' by this agricultural adaptation of the model, thus also aiding cases where predictions made by YOLOv8 may be caused to uncertainty. In addition to improving classification, VGG16 functions by looking at the high-level features contained in a specific cropped region to minimize the chances of errors occurring, especially in highly vegetated regions or when the architecture of crops and weeds overlaps. This helps to improve the accuracy of the images aimed at identifying weeds, which is essential with regard to precision agriculture.

Interpretability with Grad-CAM

In order to enhance model interpretability and help end users trust the model in the detection tasks accuracy, Grad-CAM is used to highlight the salient regions in the respective classification model. Users are thus able to appreciate and endorse the results of the model thanks to Grad-CAM transparent provision of the areas of each classification decision, which allows the model outputs to conform with common sense agricultural knowledge.

Pipeline Workflow

The procedure of the methodology flows in this order. First, CycleGAN does an image transformation to balance the data. Second, YOLOv8 detects and localizes crops and weeds. Third, the detections are refined using VGG16. Finally, visual reasoning is done via Grad-CAM. The images produced in these methods are correctly labelled and interpretable. Each component is unattached all the while, and there are different working spaces for both YOLOv8 and VGG16 for easy updates without being rigid. This approach makes strong weed detection and classification systems possible while allowing for incorporate data oriented agricultural systems for this method of management by adding data augmentation, dual model processing, and model interpretation.

This innovative strategy mixes effective weed detection with image diversification and seeks weed identification over time, hence resolve issues of weed identification in real time. Incorporation of CycleGAN, YOLOv8, VGG16 and Grad-CAM maximize the model output and understanding, providing an adaptable system for future use in crop observation and management in agriculture.

Performance Metrics

To assess the performance of the proposed weed detection system, several measures were adopted to measure how well the system detects, classifies and transforms images of a specific object. These measures give a relational account of how effective the system is, in being able to detect and classify crops and weeds accurately; as well as detail the extent synthetic images designed for improving model training were of quality.

1. **Object Detection Metrics (YOLOv8)**

Precision: Precision is determined by the true positive detections of weeds or crops, which are all the countries weeds cropped up detected by YOLOv8. This metric serves to gauge the extent to which the model is able to eliminate false positives thus determining reliable weed targeting within crops images.

Recall: Recall presents the true positive detection rate in relation to all available population in the dataset. High recall means that the model correctly finds most instances of weeds and crops in each of the pictures.

mAP (Mean Average Precision): Mean Average Precision (mAP) at 0/5 intersection of different targets is mAP at 0.5 in a mean average precision evaluation. mAP @ 0.5 is used to evaluate the overall detection performance of the detection system for all classes. This is done to provide a single measure that reflects the detection effectiveness of the model for each class by calculating the averages of precision metrics over several recall values for the class.

IoU (Intersection over Union): This is a measure of the extent to which the predicted bounding box and the actual bounding box intersect, thus indicating how well the model localizes the objects that it has detected. High IoU values imply that the crops and weeds are well bound within the image.

2. **Classification Metrics (VGG16)**

Accuracy: The classification accuracy is defined as the ratio of the number of correct classifications (for instance “carrot”, “potato” or “weed”) made by the VGG16 model to the total number of classifications done by the model. This metric assesses how well the model classifies differences of crops and weeds inside the bounding boxes bounding the detected objects.

F1-Score: The F1 score combines precision and recall to produce in effect a single level of performance for a given classification. It is especially important for VGG16 classification accuracy evaluation as it provides a trustworthy metric in classification performances when the distribution of the classes is not uniform.

Confusion Matrix: To investigate the performance of the model with respect to individual classes, a confusion matrix was constructed, indicating the frequency with which each type of crop or weed was identified correctly and incorrectly. This matrix identifies weaknesses in the classification and specifies the need for improved classification accuracy.

3. **Image Transformation Quality (CycleGAN)**

SSIM (Structural Similarity Index): This index is employed to compare original and transformed images in terms of structure, luminance, and contrast. The closer the value of SSIM is to, the more likely those images are close to the clean crop images thus the clean images that are generated using CycleGAN augmentation for training will be of good quality for the training.

PSNR (Peak Signal-to-Noise Ratio): Peak Signal to Noise Ratio is used in determining the impact of the images create by CycleGAN against the original ones. The higher the PSNR, the less the noise and artifacts present in the images, hence more suitable for transformation of images for model training.

Cycle Consistency Loss: This refers to the difference found between the input images and the images which have been obtained after a complete sense – transform out and back (e.g. mixed to clean then back to mixed) process and hence assesses CycleGAN’s ability to preserve content during alterations.

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 YOLOv8, VGG16, and CycleGAN pipeline in supporting real-time, scalable weed detection in precision agriculture is demonstrated.

Dataset Description

This research is based on a dataset that comprises photographs of main crops such as carrots and potatoes. In addition, the dataset has a mixture of crop, and crop with weed images. The dataset is equipped to handle object detection and its classification tasks and in addition synthetic data generation strategies to solve the issue of class imbalance.

1. Raw Image Collection

Potato Dataset: In this potato dataset, there are potato plant images with and without weeds. The images are also drawn with boxes on the areas where the weeds or the potatoes are for example regions supporting the use of the object detection tasks.

Carrot Dataset: The carrot dataset prior had only images of carrots with weeds and did not have enough healthy looking (free of weeds) images. To counter this limitation, clean carrot images were synthesized through a CycleGAN model from mixed carrot with weed images which improved the dataset balance as well as its performance across varying crop conditions.

2. Image Annotation

Bounding Box Annotations: Images, both of the potato and carrot datasets, have been annotated with the boxes around the crops and weeds which are present in the images. The annotation was conducted using the Visual Object Tagging Tool (VoTT) where the objects were labeled 'weed', 'carrot' or 'potato' in order to facilitate the training of the YOLOv8 model.

Class Labels for Classification: Any labelled images where two or three class labels appear at the same time. In relation to the VGG16 classification model, the images are also decomposed into hierarchically structured classes, with each class e.g. 'weed', 'carrot', and 'potato' having its training and validation folder structures created.

3. Synthetic Image Generation

CycleGAN for Dataset Augmentation: Initially, it was hard to collect plenty of clean images of carrots but later on CycleGAN was applied for the conversion of images of carrots with weeds into those of clean carrots only. This modification allowed the modification to the dataset with variety of clean examples without weeds aiding the model to generalize in different cropping conditions. The images synthesized using the CycleGAN are realistic and very closely depicted comparable to original images which have been confirmed using parameters like SSIM and PSNR for suitability in training processes.

4. Dataset Split

The dataset divided is further used in the object detection and classification tasks hence the ration also supportive of having training and testing the object detection system. The integration of object detection training processes involves the use of image crops with appropriate bounding boxes indicating where the weeds and crops are located. For classification purpose images are rather separated into training and validation folders class wise so as to facilitate the VGG16 model that finetunes the outputs of YOLOv8 on.

The dataset is created in such a way, using both annotated images of actual objects and their computer-generated models, that it minimizes the risks of the model failure in performing the different tasks which also includes the localized serving of weed detection in carrots and potatoes put in actual field conditions.

Architecture

The suggested architecture for the weed detection system is structured as a multi-stage process comprising CycleGAN, YOLOv8, VGG16, and Grad-CAM, with each participant uniquely contributing towards the attainment of the primary weed detection objective. As such, the process starts with CycleGAN which is the most initial step of the process, where images full of crops and weeds are mixed to synthesize clean images only containing crops. This increases the diversity of the dataset and also balances the classes hence solving under-represented classes. After data augmentation, the primary detection model becomes YOLOv8 that detects and locates the weeds and crops in the images by drawing boxes around the objects. All regions of interest are forwarded to VGG16 for more accurate classification to separate weeds from the crops of interest in order for the output to be precise. To the output image from VGG16, regions of interest by the model are added using Grad-CAM, which provides visualization features, thus enhancing the ability to understand the results based on how the classification was made.

Such an architecture allows for the use of data transformation capabilities within CycleGAN, object localization capabilities within YOLOv8, detail classification of various objects within VGG16 and visualization within Grad-CAM, resulting to systematized weed detection process that is both effective and clear in its purpose.

CycleGAN for Synthetic Data Generation

For that purpose, in this work, CycleGAN was used to tackle the dataset problem as well as help model generalization by creating synthetic clean images with no weeds from the carrot weed compilations. The volume of realistic "crop + weed" blended images was enlarged with the help of CycleGAN, which allowed to create "clean" images out of dirt-infected carrot-and-weed images- something which was highly critical in ensuring proper weed management in varying conditions within agricultural fields.

1. Purpose and Process

The main goal of introducing CycleGAN was to use it inversely to create clean crop images from the images that showed weeds, to help balance the dataset, such that no weeds are in look found crops. This image-to-image translation format helps the model to be trained on multiple images making it better at weed detection in different circumstances. The results of CycleGAN transformations are shown in Figure 1. The model is capable of removing weeds (Fake B), and generates images in which weeds are present from images where no weeds are seen (Fake A), which shows that both images of carrot in weedy and non-weedy conditions can be generated with ease.

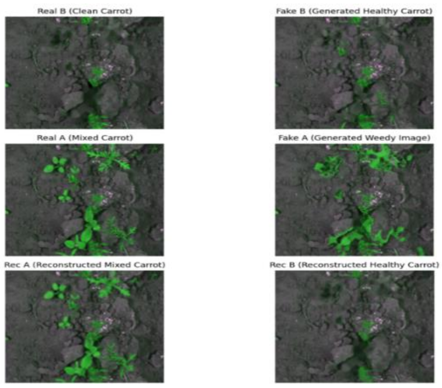

Figure 1: CycleGAN Image Translation Results

Another example of the clean image generation is given in Figure 2 where CycleGAN demonstrates its ability to create clean images devoid of weeds from inputs containing input images containing weeds. This process allows generating a dataset that suits better for the model training by ensuring that there are sufficient images for both weedy and non-weedy cases.

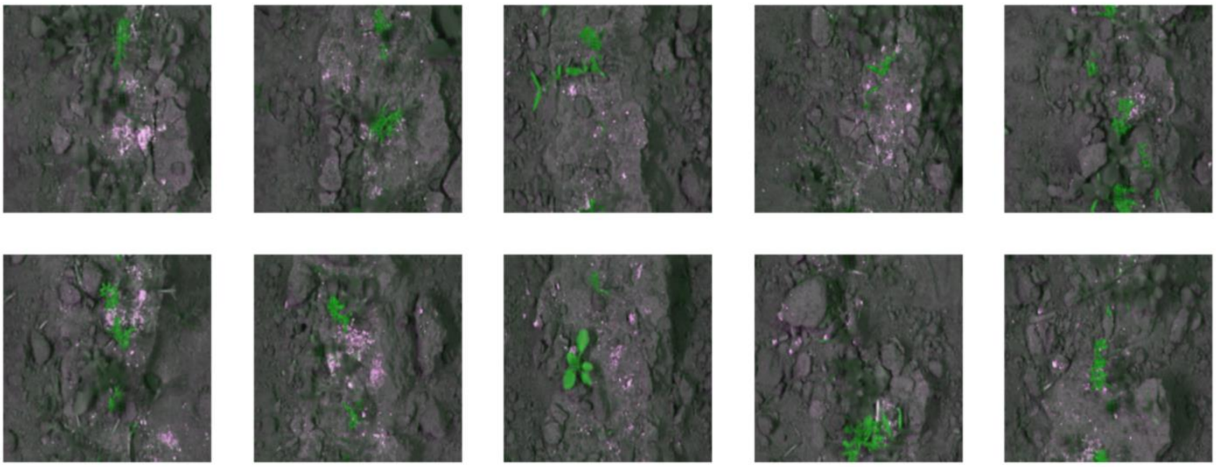

Figure 2: Weed-free carrot image generated by CycleGAN (Output).

2. Training and Loss Convergence

In particular, the training was carried out by minimizing maximal two types of losses, discriminator losses (D_A, D_B) and generator losses (G_A, G_B), with the cycle consistency loss added in to guarantee that image translations are done accurately. In Figure 3 it can be seen that both generator and discriminator D losses decrease and then remain steady over time signifying convergence and stable training. Smaller generator and discriminator losses indicate that the process of transformation has not only been learned by CycleGAN but also that the transformation does not distort the actual image. This convergence is supportive for the image quality as it promotes that clean crops can be generated using CycleGAN without over-synthesizing latent features.

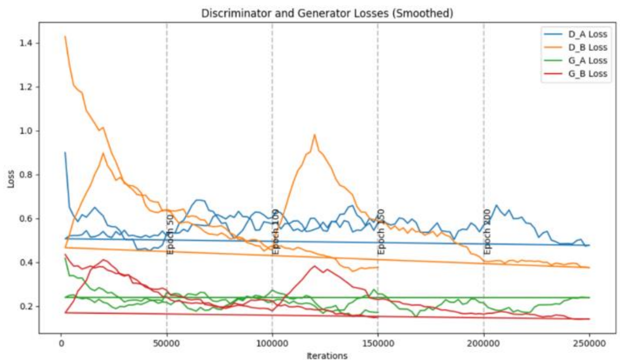

Figure 3: Discriminator and Generator Losses

As shown in Figure 4, the cycle consistency losses exhibit a downward trend with increasing training epochs demonstrating that it is possible to transform images (e.g., mixed to clean, and vice versa) while preserving certain features. This stabilization of the cycle consistency loss also allows for the content and the arrangement of the crops in bushes to be kept intact despite the change of domains which is important in synthesizing hyper-real images.

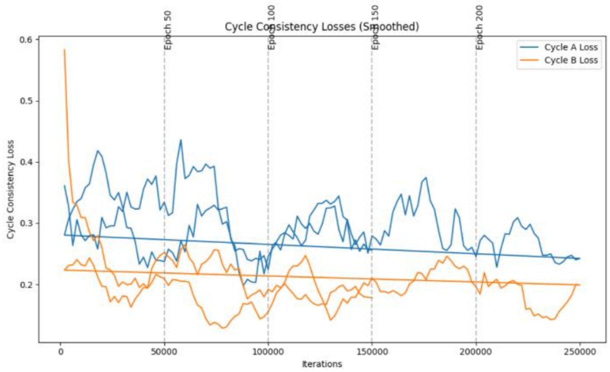

Figure 4: Cycle Consistency Losses

3. Evaluation of Generated Images with SSIM and PSNR

The quality of images created by CycleGAN was examined through the usage of SSIM (structural similarity index measures)/criteria and along with PSNR (peak signal-to-noise ratio) image ratings which are all presented in the figures figure 5 and figure 6 respectively. The SSIM - sophisticated similarity measures provided in -range 0-1 for all the test images produced images of moderate similarity to image of the actual target hence retained most if not all the structural features in the real clean images. The difference in SSIM score is also due to the fact that some images are more complicated such as the image where weed removal or weed replication would be much difficult to execute.

Similarly, PSNR values also evaluate the perceived quality of generated images. PSNR figures ranging between 18-20 dB are of a good standard for images meant for training purposes, though some types of images (like those with complicated weed patterns) drop in PSNR availability. The highest values of PSNR (Figure 6, Image Index 8 contained) are for simpler images when the model correctly recreated a clean field with noise in much smaller amounts, indicating that CycleGAN can generate high-quality images but simple ones with lesser complication levels.

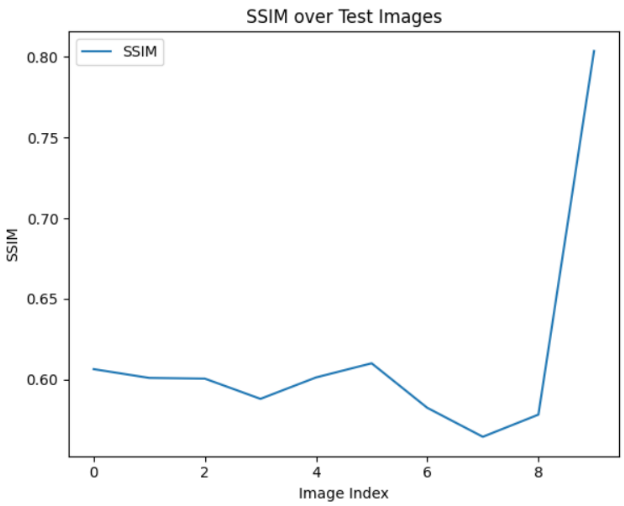

Figure 5: SSIM over Test Images

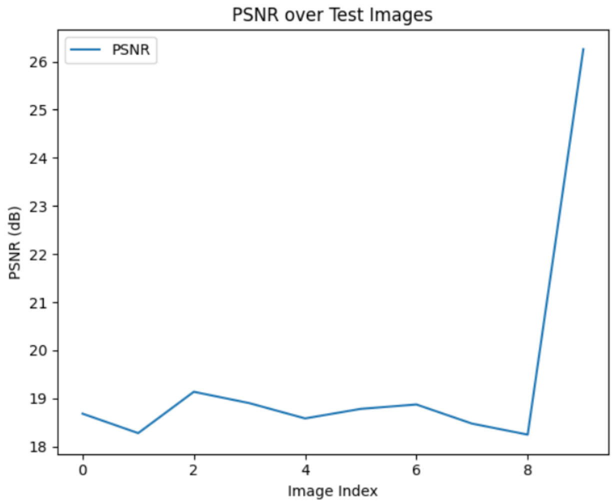

Figure 6: PSNR over Test Images

4. Impact of CycleGAN on Model Performance

The incorporation of images from CycleGAN processed images in the weed detection system was effective because of the following reasons:

- **Improving Dataset Balance**: The model managed to generalize well in the weeded and weed-free conditions and therefore minimized overfitting which in turn increased robustness.
- **Enhancing Detection Accuracy**: Synthetic images generated by CycleGAN helped the model comprehend patterns in both weeded and unweeded crop landscapes thus aiding separation of the weeds during the test process more accurately than without cycle.

The overall contribution of CycleGAN towards the production of realistic images proved to be instrumental in the weed detection process which encourages the use of such tools in the agricultural field considering such technology will require accessing very diverse datasets.

YOLOv5 for Object Detection

Mainly, the weed detection model's object detection components utilized the YOLOv8 model with particular attention on the crops and weeds (carrots and potatoes) regions detection and localization. The architecture of the YOLOv8 model is designed for speed and accuracy making it appropriate for real-time operations in precision agriculture where instant and precise detection is required.

1. Purpose and Function of YOLOv5 in Weed Detection

YOLOv8 is an integral member of the processing pipeline where it carries out its primary task of object detection in image data by generating bounding boxes on the suspected images of weeds and crops. As it can quickly find the regions of interest, the system is able to focus on the crops and weeds in every image, which can then be classified and refined in further stages. Such localization ability comes in quite handy in agricultural fields, considering the fact how difficult it is to visually separate object in instances mared by dense arrangement of overlapping plants.

2. YOLOv8 Model Setup and Training Parameters

The training and optimization of YOLOv8 model was preformed using augmented dataset that consists original and CycleGAN generated synthetic images. The following parameters were distinct:

- **Image Size**: 640x640 pixels, where resolution and computational power were well proportioned.
- **Batch Size**: 32, which enabled quicker convergence in the course of training.

- **Number of Epochs**: 50, which was done to ensure enough iterations were carried out to learn the detection of crops and weeds from complex backgrounds..
- **Evaluation Metrics**: The evaluation of the model was based on precision, recall and mean Average Precision (mAP) for different groups.

3. Detection Performance and Results

The results on the performance of the YOLOv8 algorithm on weeds, carrots, and potatoes are summarized in Table 1. These indicators reveal the efficiency of the model for these classes as well as for the performance in general.

Class	Images	Instances	Precision	Recall	mAP@50	mAP@50-95
all	41	139	0.70542	0.6604	0.71331	0.3283

Table 1: YOLOv8 Detection Performance

With an overall precision of 0.705 and a recall of 0.660, this model localizes crops and weeds well, but there is some potential for improvement, especially in differentiating between weeds and potatoes. Further, the mAP@50 score of 0.713 reveals the model's capability to detect objects rather accurately, across all classes. On the contrary, the lower mAP@50-95 of 0.3283 reveals the existence of the change in detection's accuracy at varying levels of Intersection over Union (IoU), which shows that there might be some difficulties in maintaining consistent localization across scenarios.

This performance gap implies that more model tuning or tailored data augmentation could be beneficial in improving the performance of the machine especially for classes such as weeds and potatoes where confusion is common.

4. Analysis and Discussion of Detection Results

The model was able to detect carrots with higher accuracy compared to weeds and potatoes owing to the clearer visual characteristics of carrots that are also aided in by the synthetic clear images provided by the CycleGAN. The precision and recall figures that were observed for weeds are relatively low and this indicates that perhaps more specialized data or training phase that is focused on weed samples would help improve performance in detecting weeds. Something similar can be said of the potato class as it was problematic which can be due to either lack of diversity in the data or the fact that potatoes have some background elements that are very similar to them, thus more confused the model in determining them accurately.

Nonetheless, the ability of the model to operate at a high speed and efficient performance in the detection of carrots highlights the model application in real time detection of weeds. Further, the model optimally trained for detecting underrepresented classes such as weeds and potatoes could potentially be more useful and flexible in various agricultural practices. The following steps of the pipeline employ VGG16 for further improvement such as enhancing accuracy and credibility of the weed classification system which in the end promotes precision agriculture systems.

VGG16 for Classification Refinement

The VGG16 model was integrated as an additional classification layer to refine the predictions made by YOLOv8. Under the normal operating conditions, when the images are taken and analyzed by the system, YOLOv8 performs the detection and localization of the crops and weeds, while the VGG16 processes those marks by classifying them into carrot, potato, and weed detection. As a result, the system uses VGG16 in order to compensate for the low level of confidence found in the initial output of the YOLOv8 model or where it is expected that the predictions from the model are questionable.

1. Purpose and Role of VGG16

As already mentioned, VGG16 is used here in the pipeline to address the problem of correctly classifying crops and weeds in the bounding box created by the YOLOv8 system. This additional classification step helps the system reduce the number of false zones and improves the accuracy of distinguishing crops vs weeds. Once again, the system crops some regions out from YOLOv8, passes them to VGG16, and helps to place better labels, thus enhancing overall precision-recall metrics with respect to the detection pipeline.

2. Model Training and Evaluation Metrics

The VGG16 model was further trained on the crops – weeds database, where the classification head was modified to suit carrot, potato and weeds and utilized ImageNet pretrained network for features extraction. The training process included:

- **Learning Rate**: 1×10^{-4}, appropriate for tuning the last few layers as the earlier layers were frozen.
- **Batch Size**: 32, which is adequate for performing learning without straining the memory.
- **Epochs**: 20, with use of early stopping technique to control for overfitting.

Model results are presented in Table 2, where the performance of the model class measures including precision, recall, F1-score, and support are included.

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 F1-scores for weeds and carrots namely 0.96 and 0.91 respectively, attest to the efficacy of theVGG16 model in classifying these classes as it attains high precision and recall. On a different note, the class potato has a lower performance in recall with the score of 0.80, indicating that other potatoes in the test run were either neglected or misplaced. Nonetheless, the figure of 91 percent overall accuracy demonstrates the strength of this model in differentiating between crops and weeds.

3. Confusion Matrix and Analysis

In Figure 7, the confusion matrix undergoes additional scrutiny to assess classification precision by including true and false positive and negative values for each class. As can be seen, the model perfectly classifies the carrot class without a single instance of misclassification in this category. For potatoes, as many as 8 were returned as carrots and 2 as weeds, indicating some degree of visual similarity between the two class categories. The weed class only returns a single misclassification of carrot, demonstrating that the model is able to differentiate between weeds with other classes.

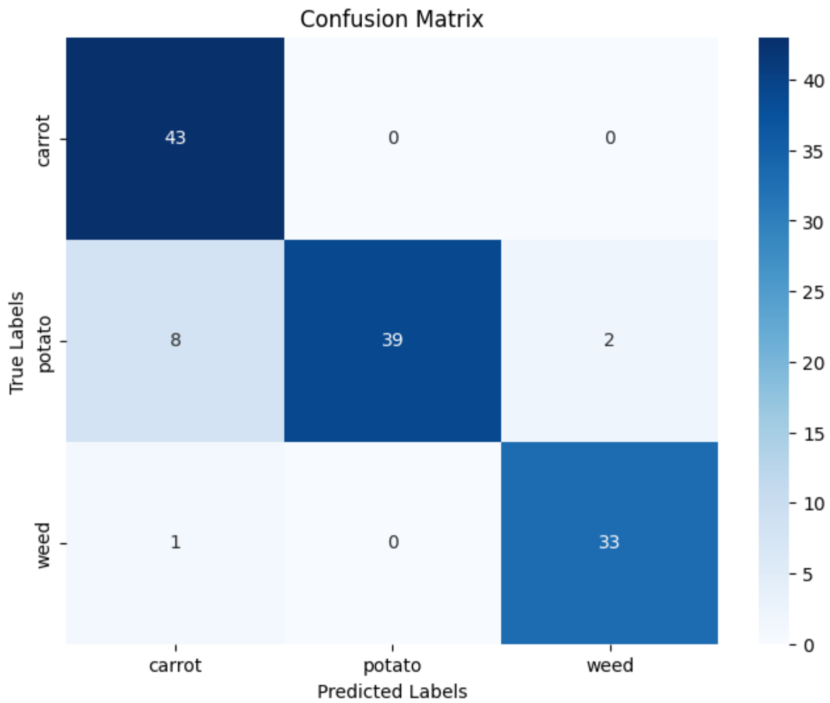

Figure 7: Confusion Matrix of VGG16 Classification Results

The impressive classification performance and little class agonistic rivalry seem to support the conclusion that the model VGG16 is a good addition to YOLOv8 and it helps in minimising the instances of sitting crops and weeds, in turn, improving the overall performance of the weed detection system.

4. Grad-CAM for Model Interpretability

With a view to improving transparency and making the users more confident about the model, in this case Grad-CAM is utilized on the top of VGG16 its output in order to draw heatmaps showing where most focus is being solicited for every classification made. Thanks to this interpretability layer the user can see where VGG16 looks at while classifying crops and weeds, therefore improving the trust level to the model's outputs and making sure that the classifications are consistent with what is found in fields.

5. Discussion of VGG16's Impact on the Pipeline

Deployment of VGG16 in the pipeline can be seen to enhance the classification accuracy to a greater extent even for those images which the predictions of the system using YOLOv8 may be somewhat ambiguous. Instead of redirecting the false alarms according to the primary detection of Yolov8, which cannot be avoided altogether, VGG16 reinforces the scope of crops and weeds eliminated the interference of other objects. In addition, Grad-CAM brings a break up and explanation of how the process of classifying has been put into practice which then helps in instilling confidence in the model's outcomes to the end-users. The fact that there is a pre-trained model that has been adapted to this specific field of agriculture dataset, allows for a practical application without having to go through a lengthy training process. Hence, this makes VGG16 an asset in the entire detection architecture.

Integration for Application Use

The integrated weed detection pipeline incorporates CycleGAN, YOLOv8, VGG16, and Grad-CAM within a modular framework which is application-friendly to precision agriculture for its real-time aspect. CycleGAN in the beginning creates synthetic clean images for dataset balancing, a technique which increase the performance of YOLOv8 and VGG16 under different crop habitats. Next, YOLOv8 does object detection, detects boxes on the image, after which VGG16 provides an even more refined regions of interest which classifies the area as crops or weeds. Finally, Grad-CAM examples are applied to the images to show which areas of the image were focused on by the model, such that the field workers can confirm the accuracy of detections.

For application use, this integrated system may be fitted on agricultural drones or installed on a phone in the field. With the rapid detection arms of YOLOv8, the fine-tune ability of VGG16, and the clarity of purpose afforded by Grad-CAM, the system allows for fast and effective weed management while targeting the specific problem area. Because this solution enables the sequential integration of all its components into a workflow, it assists farmers in enhanced identification and monitoring of weed infestations, lessening the dependence on herbicides and increasing the yield of crops.

Conclusion

This paper employed different kinds of algorithms which include CycleGAN, YOLOv8, VGG16, and Grad-CAM in developing an integrated approach to weed detection. The system was able to detect weeds in crops more effectively by using CycleGAN for synthetic data creation, adjusting to the object detection task with easy YOLOv8 and further classification using VGG16 network with Grad-CAM for attributed visualization of the results. The CycleGAN component, for instance, solved the problem of insufficiency of data by creating clean background images devoid of weeds hence establishing a well-balanced training data for enhanced model performance. Furthermore, YOLOv8 was efficient in locating both the crops and weeds present where VGG16 came in to solve the problem of more and less accurate classifications to eliminate the false positives of crops and weeds and classified them very well. Moreover, Grad-CAM provided a nice addition to the system by providing visual illustrations of the areas targeted by the model which increased the trust levels of the output provided by the model. This pipeline provides a deployable architecture for applications ranging from the real world small scale systems to the field and contributes towards sustainable agriculture by means of efficient management of weeds and less reliance on the herbicides..

Future Work

More attention could be given to the improvement of the model's capabilities in relation to other crops and environmental conditions in which farming is carried out. The use of other sensors, for example, hyperspectral or multispectral imaging, may enhance the ability to dissociate the weeds by spectral data obtained from the plant surface. Increasing the database with weeds and crops at different growth stages would help increase the generalization and precision of the model. Other sophisticated deep learning models like attention-based networks can be explored which would boost the detection process in other complicated cases. Finally, integrated systems of such kind with reduced and optimized processing pipelines on edge devices are possible which would allow real time "weed and spill" detection and control over a wider area and thus making the solution more helpful and feasible for the farmers across the globe.

Reference

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