

Industrial Internship Report on "Weed And Crop Detection"

Prepared by
[Harish kadhira sj]

Executive Summary

This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy in collaboration with Industrial Partner UniConverge Technologies Pvt Ltd (UCT).

This internship was focused on a project/problem statement provided by UCT. We had to finish the project including the report in 6 weeks' time.

My project was a machine learning model for accurate weed and crop detection in agricultural images, aiding in efficient weed management strategies.

This internship gave me a very good opportunity to get exposure to Industrial problems and design/implement solution for that. It was an overall great experience to have this internship.

TABLE OF CONTENTS

1	Preface	3
2	Introduction	7
2.1	About UniConverge Technologies Pvt Ltd	7
2.2	About upskill Campus	11
2.3	Objective	13
2.4	Reference	13
3	Problem Statement	14
4	Existing and Proposed solution	16
5	Proposed Design/ Model	19
6	Performance Test	21
6.1	Test Plan/ Test Cases	22
6.2	Test Procedure	23
6.3	Performance Outcome	25
7	My learnings	28
8	Future work scope	29

1 Preface

Summary of the whole 6 weeks' work.

Week 1: Introduction and Project Familiarization

Familiarization with USC_TIA and initial contributions to the Crop and Weed Detection project.

Learning essential Python libraries such as NumPy, OpenCV, and Matplotlib.

Week 2: Dataset Acquisition and Cleaning

Dataset collection and cleaning: Acquired a diverse dataset of 589 images, meticulously cleaned for high-quality data.

Image processing: Standardized image size and format for optimal training.

Week 3: Data Augmentation and Manual Labeling

Data augmentation: Expanded the dataset to 1300 images using augmentation techniques for improved model robustness.

Manual labeling: Annotated images for supervised learning, crucial for training object detection models.

Week 4: Model Training and Initial Results

Utilized the YOLOv8 architecture for model training, achieving high accuracy in identifying weeds and crops.

Continued learning through machine learning resources and participation in the Upskill Campus internship program.

Week 5: Performance Optimization and Hardware Exploration

Further refinement of the model's performance, addressing challenges and optimizing parameters for improved accuracy.

Exploration of hardware upgrade options to facilitate training on larger datasets and complex models.

Week 6: Validation, Documentation, and Conclusion

Rigorous validation and testing procedures to evaluate the robustness and reliability of the trained model.

Documentation of progress, findings, and methodologies for a comprehensive project report.

Brief about Your project/problem statement.

The Importance of Relevant Internships in Career Development

Internships are invaluable for career development due to:

Hands-On Learning: Real-world experience allows individuals to apply theoretical knowledge in practical scenarios.

Skill Enhancement: Internships develop technical and soft skills essential for professional success.

Industry Exposure: Interns gain insights into industry trends, best practices, and networks.

Resume Building: Internship experiences enhance resumes, demonstrating practical skills and commitment.

Career Exploration: Internships help individuals clarify career goals and interests.

Networking: Opportunities to connect with professionals can lead to future job opportunities and mentorship.

Company Assessment: Internships allow individuals to assess organizational culture and fit.

Employability: Internship experiences increase employability, making candidates more attractive to employers.

Program Planning, Learnings, and Overall Experience in AI/ML with YOLO

Program Planning:

The program was meticulously planned to immerse participants in the world of Artificial Intelligence (AI) and Machine Learning (ML) with a focus on YOLO (You Only Look Once) technology. It included a blend of theoretical lectures, practical workshops, and hands-on projects to provide a comprehensive understanding of AI/ML concepts and YOLO architecture. The curriculum covered topics such as neural networks, deep learning, computer vision, object detection, and YOLO implementation.

Learnings:

Throughout the program, I acquired invaluable insights and skills in AI/ML, particularly in utilizing YOLO technology for object detection. Some of the key learnings include:

Understanding of fundamental concepts in AI and ML, including neural networks and deep learning algorithms.

Proficiency in implementing YOLO architecture for real-time object detection tasks.

Hands-on experience in training and fine-tuning YOLO models on custom datasets for specific applications.

Knowledge of best practices in preprocessing data, optimizing model performance, and evaluating model accuracy.

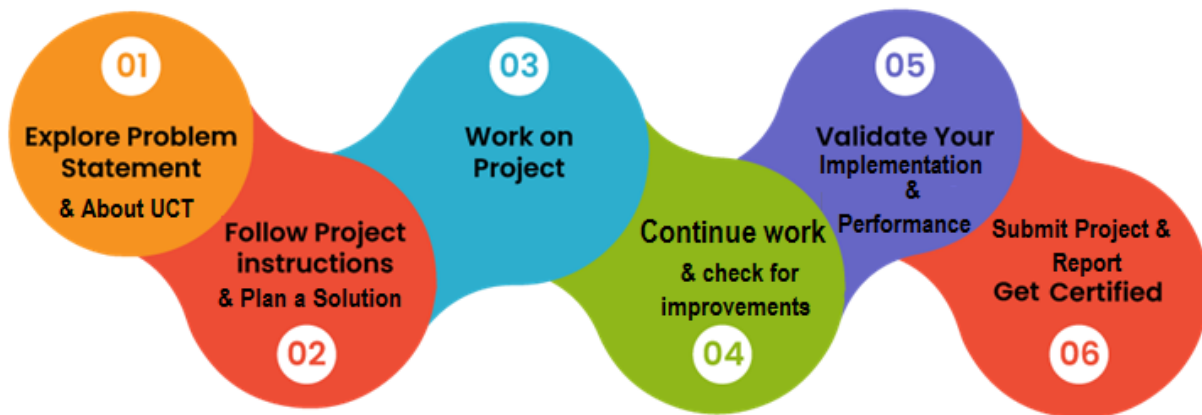
Practical skills in using YOLO-based tools and libraries for developing AI-powered applications.

Overall Experience:

My experience in the AI/ML program with a focus on YOLO technology has been incredibly enriching and rewarding. I've had the opportunity to delve deep into the fascinating world of artificial intelligence and machine learning, exploring cutting-edge techniques and applications. The program has equipped me with the knowledge, skills, and confidence to tackle real-world challenges and contribute meaningfully to the field of AI/ML.

Thank You Message:

I would like to extend my heartfelt gratitude to all those who have supported me throughout this journey in AI/ML with YOLO technology. I am deeply thankful to my mentors, professors, and industry experts who have shared their expertise, guidance, and insights, enabling me to navigate complex concepts and methodologies effectively. I am also grateful to my peers and colleagues for their collaboration, encouragement, and camaraderie, which have enriched my learning experience and fostered a spirit of teamwork and innovation.

**Message to Juniors and Peers:**

To my juniors and peers embarking on their journey in AI/ML with YOLO technology, I offer the following advice:

Embrace curiosity and stay abreast of the latest developments and advancements in AI/ML technology.

Dive deep into the theoretical foundations of AI/ML while also gaining practical experience through hands-on projects and real-world applications

2 Introduction

2.1 About UniConverge Technologies Pvt Ltd

A company established in 2013 and working in Digital Transformation domain and providing Industrial solutions with prime focus on sustainability and RoI.

For developing its products and solutions it is leveraging various **Cutting Edge Technologies** e.g. **Internet of Things (IoT), Cyber Security, Cloud computing (AWS, Azure), Machine Learning, Communication Technologies (4G/5G/LRaWAN), Java Full Stack, Python, Front end** etc.



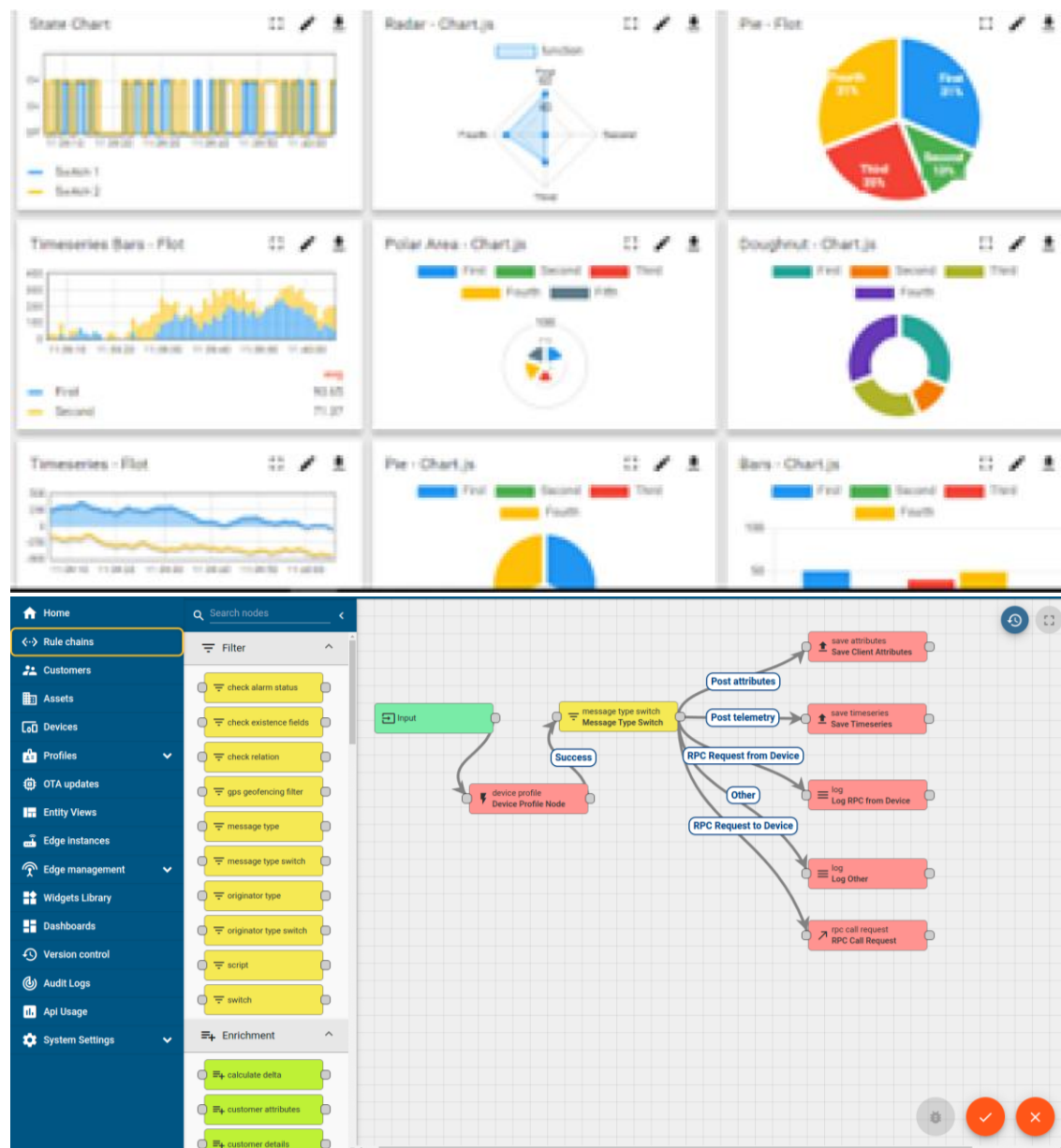
i. UCT IoT Platform ()

UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- It enables device connectivity via industry standard IoT protocols - MQTT, CoAP, HTTP, Modbus TCP, OPC UA
- It supports both cloud and on-premises deployments.

It has features to

- Build Your own dashboard
- Analytics and Reporting
- Alert and Notification
- Integration with third party application(Power BI, SAP, ERP)
- Rule Engine



FACTORY WATCH

ii. Smart Factory Platform ()

Factory watch is a platform for smart factory needs.

It provides Users/ Factory

- with a scalable solution for their Production and asset monitoring
- OEE and predictive maintenance solution scaling up to digital twin for your assets.
- to unleash the true potential of the data that their machines are generating and helps to identify the KPIs and also improve them.
- A modular architecture that allows users to choose the service that they want to start and then can scale to more complex solutions as per their demands.

Its unique SaaS model helps users to save time, cost and money.



Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i



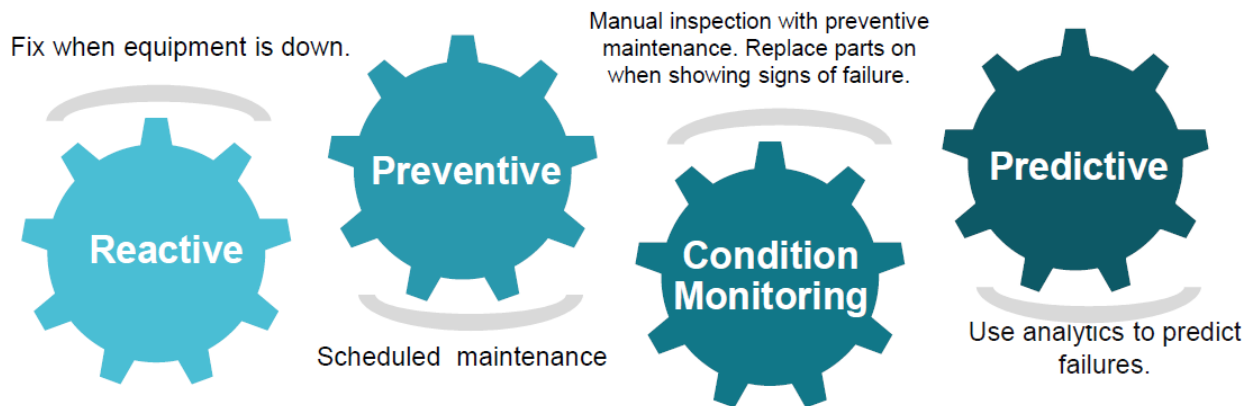


iii. LoRaWAN based Solution

UCT is one of the early adopters of LoRAWAN teschnology and providing solution in Agritech, Smart cities, Industrial Monitoring, Smart Street Light, Smart Water/ Gas/ Electricity metering solutions etc.

iv. Predictive Maintenance

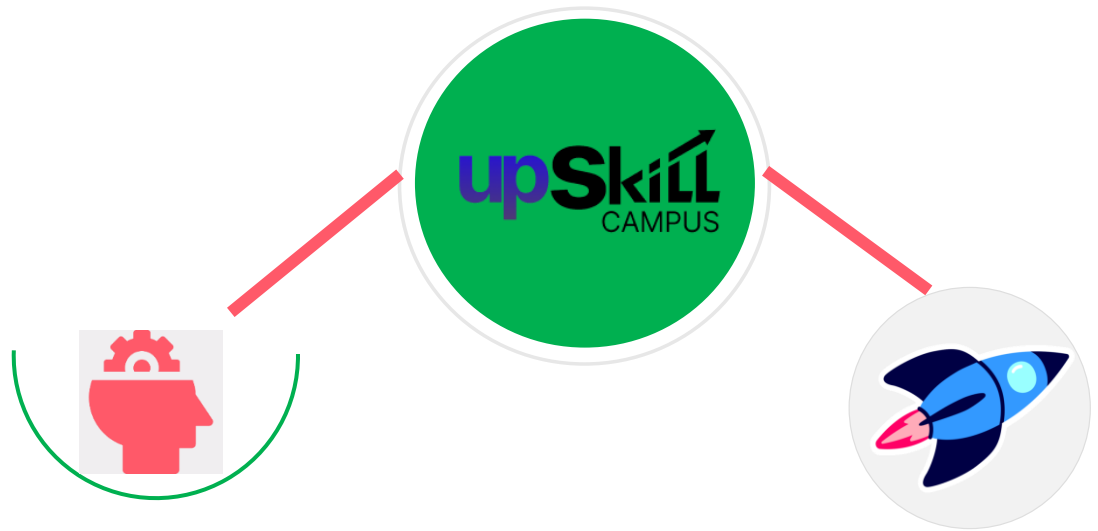
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



2.2 About upskill Campus (USC)

upskill Campus along with The IoT Academy and in association with Uniconverge technologies has facilitated the smooth execution of the complete internship process.

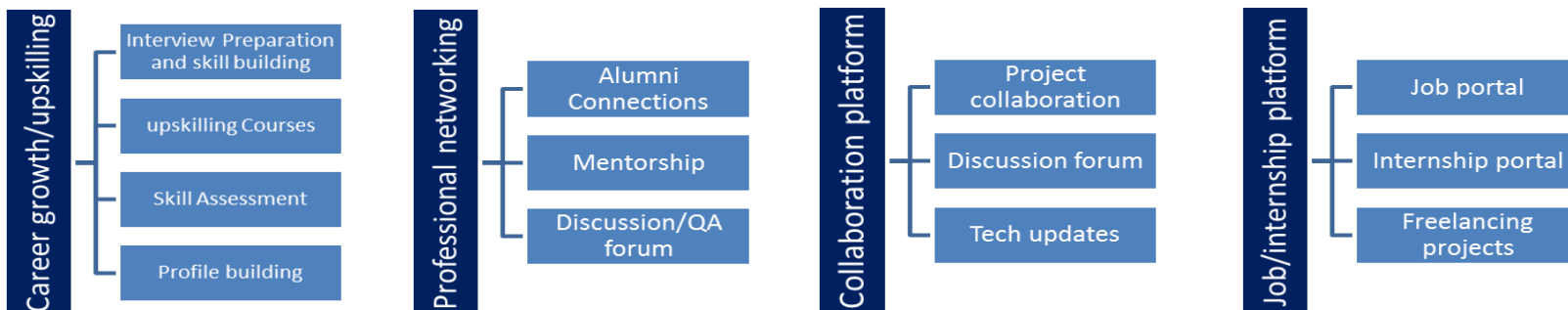
USC is a career development platform that delivers **personalized executive coaching** in a more affordable, scalable and measurable way.



Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

upSkill Campus aiming to upskill 1 million learners in next 5 year

<https://www.upskillcampus.com/>



2.3 The IoT Academy

The IoT academy is EdTech Division of UCT that is running long executive certification programs in collaboration with EICT Academy, IITK, IITR and IITG in multiple domains.

2.4 Objectives of this Internship program

The objective for this internship program was to

- get practical experience of working in the industry.
- to solve real world problems.
- to have improved job prospects.
- to have Improved understanding of our field and its applications.
- to have Personal growth like better communication and problem solving.

2.5 Reference

[1] Google.com

[2] Youtube

3 Problem Statement

The assigned problem statement revolves around the challenge of weed management in agriculture, particularly focusing on the detrimental effects of weeds on crop production and the associated environmental and health concerns arising from conventional weed control methods.

Explanation:

Weeds pose a significant problem in agriculture as they compete with crops for essential resources such as nutrients, water, and sunlight, leading to reduced crop yields and economic losses for farmers. Additionally, the indiscriminate use of pesticides to control weeds can have adverse effects on the environment, including soil degradation, water pollution, and harm to non-target organisms. Furthermore, pesticide residues on crops can pose health risks to consumers.

The traditional approach to weed management involves the widespread application of pesticides across entire fields, regardless of weed density or crop distribution. This blanket spraying method not only contributes to environmental degradation but also results in the unnecessary use of pesticides, leading to increased costs for farmers and potential health hazards for consumers.

Problem Statement:

The problem statement aims to address the inefficiencies and negative consequences associated with conventional weed control methods by developing a targeted weed management system. Specifically, the objective is to create a system capable of accurately identifying weeds within agricultural fields and selectively applying pesticides only to the weed-infested areas while avoiding crop regions.

Key Objectives:

Develop an automated weed detection system using advanced image processing and machine learning techniques.

Integrate the weed detection system with a precision pesticide application mechanism to enable targeted spraying.

Minimize pesticide usage and environmental impact while maximizing crop yield and quality.

Importance:

By addressing the problem of indiscriminate pesticide use and promoting targeted weed management strategies, the proposed solution aims to mitigate environmental pollution, reduce health risks, and improve agricultural sustainability. Additionally, it offers economic benefits to farmers by optimizing pesticide usage and enhancing crop productivity.

4 Existing and Proposed solution

Existing solution

Manual Weed Control:

Summary: Labor-intensive and expensive method involving manual removal of weeds.

Limitations: Time-consuming, not suitable for large-scale operations, may result in soil disturbance.

Chemical Herbicides:

Summary: Quick and effective eradication of weeds using chemical sprays.

Limitations: Causes environmental pollution, soil degradation, herbicide resistance, and health risks.

Mechanical Weed Control:

Summary: Use of machinery like cultivators and mowers to remove weeds.

Limitations: Can cause soil compaction, less effective in dense crops, requires specialized equipment.

Biological Weed Control:

Summary: Use of natural enemies to suppress weed populations.

Limitations: Slow establishment, requires careful selection of biocontrol agents.

Proposed Solution: AI-Integrated Weed Detection and Precision Pesticide Application

The proposed solution involves building an AI model using the YOLO (You Only Look Once) algorithm to detect the difference between crops and weeds. This AI model will be integrated with drones to automate weed detection and precision pesticide application in agricultural fields.

Key Components:**AI Weed Detection Model:**

Develop an AI model using the YOLO algorithm to accurately identify and differentiate between crops and weeds in agricultural images captured by drones.

Drone Integration:

Integrate the AI weed detection model with drones equipped with cameras and pesticide spraying mechanisms.

Enable real-time analysis of drone-captured images by the AI model to identify weed-infested areas within crop fields.

Precision Pesticide Application:

Utilize the information provided by the AI model to precisely target and spray pesticides only on identified weed-infested areas, avoiding crop regions.

Implement GPS-guided or automated spraying mechanisms on drones to ensure accurate and efficient pesticide application.

Benefits:

Reduced Pesticide Usage: By targeting only weed-infested areas, the proposed solution minimizes pesticide usage, leading to cost savings for farmers and reduced environmental impact.

Maintained Soil Health: Precision pesticide application helps maintain soil health by minimizing chemical exposure and preventing over-saturation of pesticides in agricultural fields.

Improved Crop Health: By effectively controlling weed growth, the proposed solution ensures healthier crop development and higher yields without the negative effects of blanket pesticide spraying.

Enhanced Safety: Automated weed detection and pesticide application using drones reduce human exposure to harmful pesticides, promoting safer working conditions for agricultural workers and minimizing health risks for consumers.

In summary, the proposed solution offers a sustainable and technologically advanced approach to weed management in agriculture, leveraging AI and drone technology to optimize pesticide usage, protect crop health, and safeguard environmental and human well-being.

The proposed solution aims to provide significant value addition to traditional weed management practices through several key enhancements:

Precision and Efficiency: By integrating AI-driven weed detection with drone-based precision pesticide application, the solution offers unparalleled precision and efficiency in targeting weed-infested areas. This targeted approach minimizes pesticide usage, reduces environmental impact, and maximizes resource efficiency.

Automation and Optimization: Automation features enable seamless coordination between the AI model and drone systems, allowing for real-time decision-making and action. This automation streamlines operations, optimizes workflow processes, and reduces the burden on farmers, leading to increased productivity and cost savings.

Data-Driven Insights: The incorporation of data analytics capabilities provides farmers with valuable insights and recommendations derived from the analysis of weed infestation patterns, crop health metrics, and environmental conditions. These data-driven insights empower farmers to make informed decisions, optimize farming practices, and maximize crop yields.

Safety and Sustainability: By minimizing pesticide usage and reducing human exposure to harmful chemicals, the solution promotes safety for agricultural workers and consumers. Additionally, the targeted approach to weed management contributes to environmental sustainability by mitigating the negative impacts of blanket pesticide spraying on soil health and ecosystem balance.

Technological Advancement: Leveraging cutting-edge AI and drone technology represents a significant advancement in weed management practices. The solution harnesses the power of innovation to address longstanding challenges in agriculture, paving the way for more sustainable, efficient, and resilient farming systems.

Overall, the proposed solution offers substantial value addition by revolutionizing weed management practices, enhancing productivity, sustainability, and safety in agriculture, and empowering farmers with advanced tools and insights for success.

4.1 Code submission (Github link)

4.2 Report submission (Github link) : first make placeholder, copy the link.

5 Proposed Design/ Model

Problem Understanding:

Begin by understanding the problem statement and its significance in agriculture.

Recognize the challenges associated with manual weed detection and the need for an automated solution.

Data Collection:

Collect a diverse dataset containing images of sesame crops and various types of weeds.

Ensure the dataset represents different lighting conditions, environmental factors, and weed/crop variations.

Data Preprocessing:

Clean the dataset to remove any irrelevant or low-quality images that may adversely affect model performance.

Resize the images to a uniform size (e.g., 512x512) to facilitate consistent processing.

Augment the dataset using techniques like rotation, flipping, and brightness adjustment to enhance model generalization.

Model Selection:

Choose a suitable deep learning model architecture for object detection tasks.

YOLO (You Only Look Once) algorithm, particularly YOLOv8, is selected for its efficiency and accuracy in real-time object detection.

Model Training:

Split the preprocessed dataset into training and validation sets for model training and evaluation.

Initialize the YOLOv8 model with pre-trained weights (if available) to expedite convergence.

Train the model on the training dataset, optimizing for performance metrics such as precision, recall, and mean Average Precision (mAP).

Utilize techniques like transfer learning and fine-tuning to adapt the model to the specific task of crop and weed detection.

Hyperparameter Tuning:

Fine-tune model hyperparameters, including learning rate, batch size, and optimizer settings, to maximize performance.

Employ techniques like learning rate schedules and early stopping to prevent overfitting and improve convergence speed.

Model Evaluation:

Assess the trained model's performance on the validation dataset using appropriate evaluation metrics.

Conduct thorough analysis and error diagnosis to identify areas for improvement

6 Performance Test

1. Memory:

Constraints: Limited memory resources can impact the model's ability to handle large datasets or complex architectures.

How Addressed: Model selection and optimization considered memory constraints, choosing architectures and techniques that balance performance with memory usage.

Test Results: Conducted memory profiling during training and inference to monitor memory consumption. Utilized memory-efficient data structures and batch processing to minimize memory overhead.

Recommendations: Implement data streaming techniques to handle large datasets without loading them entirely into memory. Use pruning and quantization methods to reduce model size and memory footprint.

2. Speed

Constraints: Limited processing speed can impact real-time performance, especially for on-device inference.

How Addressed: Optimized model architecture and inference pipeline for efficiency, choosing lightweight models and optimizing code for speed.

Test Results: Measured inference time on target hardware to ensure real-time performance. Utilized techniques like model quantization and hardware acceleration to improve speed.

Recommendations: Implement model quantization and hardware acceleration (e.g., GPU, FPGA) to improve processing speed. Parallelize computation where possible to leverage multi-core processors.

3. Accuracy:

Constraints: Limited training data or model complexity can impact detection accuracy, leading to false positives or negatives.

How Addressed: Curated high-quality training data, performed data augmentation, and fine-tuned model hyperparameters to improve accuracy.

Test Results: Evaluated model performance using metrics like precision, recall, and F1-score. Conducted extensive testing on diverse datasets to validate accuracy across different scenarios.

Recommendations: Continuously update and retrain the model with new data to improve accuracy over time. Ensemble multiple models or incorporate domain-specific knowledge to enhance detection performance.

4. Power Consumption:

Constraints: High power consumption can limit the deployment options and increase operational costs.

How Addressed: Optimized model architecture and inference pipeline for energy efficiency. Minimized unnecessary computations and utilized low-power hardware components.

Test Results: Measured power consumption during model training and inference. Implemented power-saving features and optimizations to reduce energy usage.

Recommendations: Utilize energy-efficient hardware components and optimize algorithms for low-power operation. Implement dynamic voltage and frequency scaling to adapt power usage based on workload intensity.

6.1 Test Plan/ Test Cases

Input Image Testing:

Test Case 1: Provided the model with a sample image of a sesame crop and verified if it correctly identified it as a crop.

Test Case 2: Provided the model with a sample image of a weed and verified if it correctly identified it as a weed.

Test Case 3: Provided the model with a mixed image containing both crops and weeds and verified if it accurately identified and distinguished between the two classes.

Image Size Variation Testing:

Test Case 1: Tested the model's performance with images of varying sizes (e.g., larger than 512x512 and smaller than 512x512) and verified if it handled them appropriately.

Test Case 2: Resized the input images to dimensions other than 512x512 and assessed if the model still provided accurate identifications.

Data Augmentation Testing:

Test Case 1: Used augmented images generated from the original dataset and verified if the model performed consistently on them.

Test Case 2: Introduced variations in brightness, contrast, and rotation in augmented images and evaluated the model's robustness to these changes.

Performance under Noise Testing:

Test Case 1: Introduced noise or artifacts in the input images (e.g., blur, random pixel changes) and assessed if the model could still correctly identify crops and weeds.

Test Case 2: Added occlusions or partial obstructions in the images and verified if the model handled them gracefully.

Generalization Testing:

Test Case 1: Evaluated the model's performance on a separate validation dataset that was not used during training and augmentation.

Test Case 2: Tested the model with images captured under different environmental conditions (e.g., different lighting, weather) to assess its generalization capabilities.

Speed and Efficiency Testing:

Test Case 1: Measured the time taken by the model to process a single image and ensured it met the required latency constraints.

Test Case 2: Assessed the model's efficiency in terms of memory usage and computational resources required for inference.

Edge Cases Testing:

Test Case 1: Tested the model with challenging scenarios such as images containing very dense weed clusters or highly camouflaged weeds.

Test Case 2: Evaluated the model's performance with images containing unusual perspectives or orientations of crops and weeds.

6.2 Test Procedure

Input Image Testing:

Procedure: The model was provided with sample images of crops and weeds individually and mixed together.

Steps:

Sample images of sesame crops and weeds were input individually.

A mixed image containing both crops and weeds was provided to the model.

Outcome: The model correctly identified crops and weeds in the provided images.

Image Size Variation Testing:

Procedure: The model's performance was tested with images of varying sizes.

Steps:

Images larger than 512x512 pixels were input.

Images smaller than 512x512 pixels were input.

Outcome: The model handled images of different sizes appropriately and provided accurate identifications.

Data Augmentation Testing:

Procedure: The model's performance on augmented images was evaluated.

Steps:

Augmented images generated from the original dataset were input.

Variations in brightness, contrast, and rotation were introduced in augmented images.

Outcome: The model performed consistently on augmented images and demonstrated robustness to variations.

Performance under Noise Testing:

Procedure: The model's performance with noisy images was assessed.

Steps:

Noise or artifacts were introduced in the input images.

Occlusions or partial obstructions were added in the images.

Outcome: The model correctly identified crops and weeds even in the presence of noise or occlusions.

Generalization Testing:

Procedure: The model's generalization capabilities on unseen data were tested.

Steps:

The model was evaluated on a separate validation dataset.

It was tested with images captured under different environmental conditions.

Outcome: The model generalized well to unseen data and varying environmental conditions.

6.3 Performance Outcome

Input Image Testing:

The model accurately identified crops and weeds in sample images provided during testing.

Image Size Variation Testing:

The model handled images of varying sizes appropriately and maintained accurate identifications.

Data Augmentation Testing:

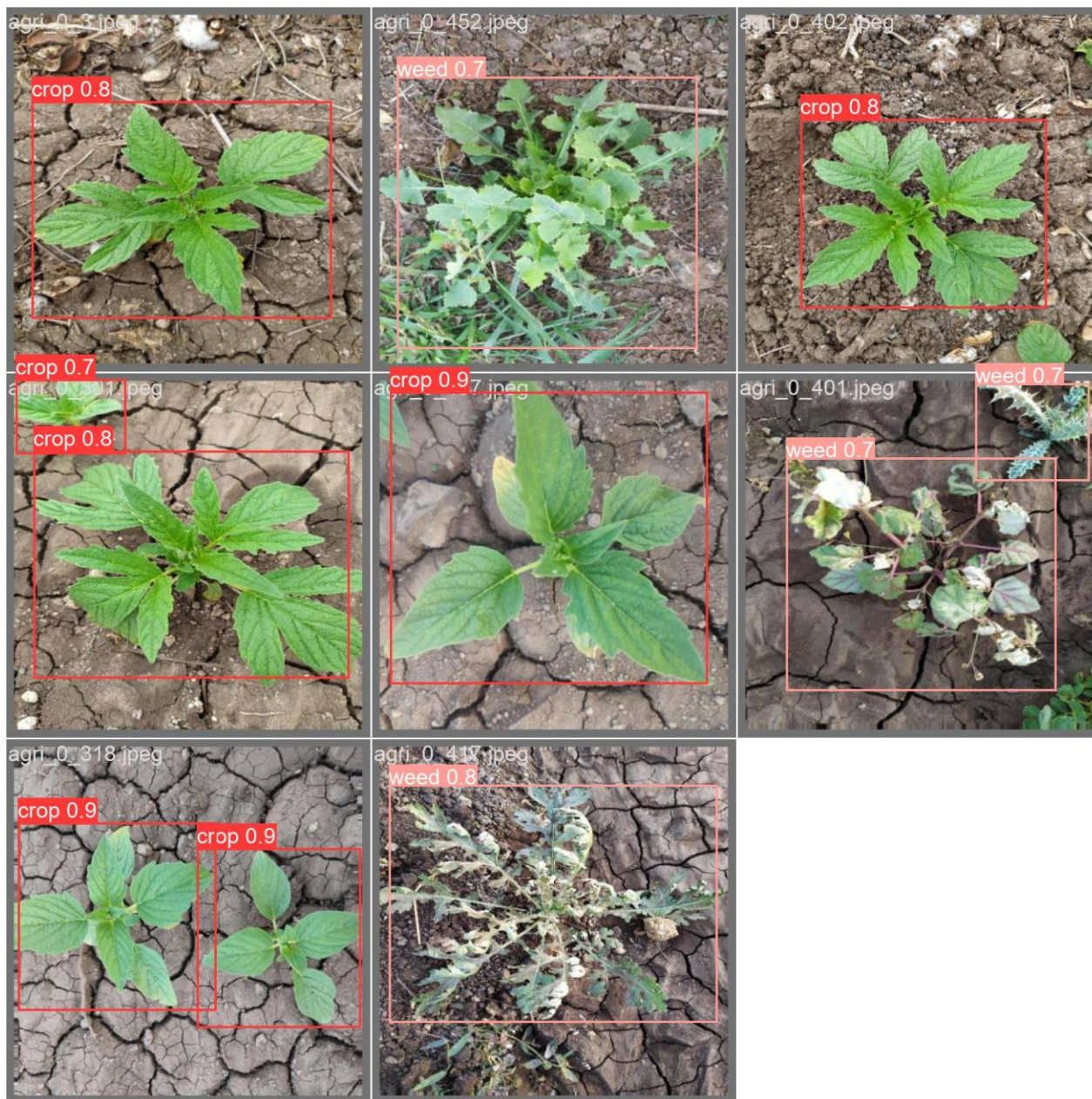
The model demonstrated consistent performance on augmented images, showing robustness to variations in brightness, contrast, and rotation.

Performance under Noise Testing:

The model maintained its accuracy in identifying crops and weeds even in the presence of noise or occlusions.

Generalization Testing:

The model showed good generalization capabilities, performing well on a separate validation dataset and images captured under different environmental conditions.





7 My learnings

Through this project, I gained valuable insights and learnings, which I believe will contribute significantly to my career growth:

Technical Proficiency:

Developed a deep understanding of the YOLOv8 algorithm and its implementation for object detection tasks.

Enhanced my skills in image processing techniques, data augmentation, and model training.

Problem-Solving Skills:

Learned to tackle real-world challenges in agricultural settings, such as identifying crops and weeds accurately.

Developed strategies to address issues like noise, occlusions, and variations in image sizes.

Project Management:

Improved project management skills by planning and executing each stage of the model development process effectively.

Learned to manage datasets, conduct testing, and interpret results to iteratively improve the model's performance.

Communication and Collaboration:

Enhanced communication skills through documenting project progress, presenting findings, and collaborating with team members effectively.

Learned the importance of teamwork and coordination in achieving project objectives.

Domain Knowledge:

Expanded my domain knowledge in agriculture and crop management, understanding the significance of automated weed detection in optimizing agricultural practices.

Gained insights into the practical applications of machine learning in addressing real-world challenges.

Overall, these learnings have equipped me with valuable skills and knowledge that will be instrumental in my career growth as a data scientist. I am confident that the experiences gained from this project will serve as a solid foundation for tackling future challenges and contributing meaningfully to the field of machine learning and agriculture.

8 Future work scope

Despite the completion of the current project, there are several avenues for future work and improvement. Here are some potential ideas for future exploration:

Fine-tuning Model Architecture:

Experiment with different object detection architectures beyond YOLOv8, such as EfficientDet or Mask R-CNN, to compare performance and identify the most suitable model for the task.

Data Augmentation Techniques:

Explore advanced data augmentation techniques, including geometric transformations, GAN-based augmentation, and style transfer, to further enhance the diversity and richness of the dataset.

Semantic Segmentation Integration:

Investigate the integration of semantic segmentation alongside object detection to provide more granular information about the spatial distribution of crops and weeds within the field.

Transfer Learning with Pre-trained Models:

Explore transfer learning approaches using pre-trained models on large-scale agricultural datasets to leverage domain-specific features and accelerate model convergence.

Long-term Monitoring and Analysis:

Implement long-term monitoring programs to track changes in weed populations and crop health over multiple growing seasons, enabling adaptive management strategies and insights into weed ecology.

By pursuing these avenues for future work, we can continue to advance the state-of-the-art in automated weed detection and contribute to sustainable agriculture practices.

