Industrial Internship Report on

”Crop-Weed Detection ”

Prepared by

Anushka Behere

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 Crop-Weed Detection, a YOLO v8 model was trained with 1300 images of crops and weeds over 20 Epochs. Testing images were then accurately predicted with an impressive 82% accuracy rate, showcasing the effectiveness of the trained model in identifying and distinguishing between crops and weeds. The project demonstrated notable success in employing machine learning techniques for precise identification and differentiation between crops and weeds, crucial for agricultural applications. 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

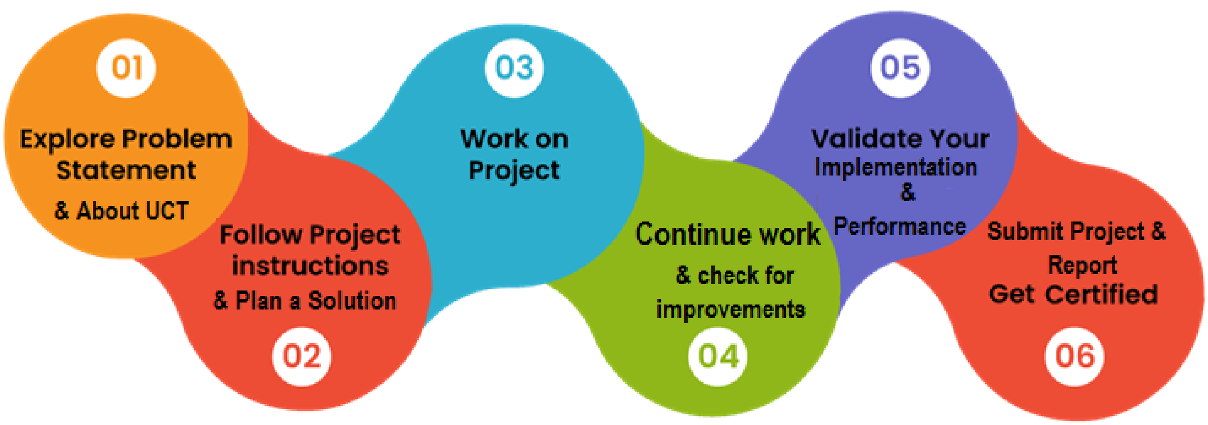
Over the six weeks of my internship, I've had an incredible learning journey. Internships are super important for anyone's career. This internship experience has not only deepened my understanding of complex concepts but also equipped me with invaluable skills essential for my career development in the rapidly evolving tech industry.

The relevance of securing a meaningful internship cannot be overstated in today's competitive job market. It serves as a crucial stepping stone for students and professionals alike, providing hands-on experience, exposure to real-world challenges, and networking opportunities crucial for career growth.

My main project was "Crop-Weed Detection", it was about using technology to help farmers. We wanted to make a program that could tell the difference between crops and weeds. This is really important for farming because it helps farmers grow more food.

I'm thankful to USC/UCT for giving me this chance to work on such an interesting project. It was a great opportunity to apply what I've learned in school to a real-life problem.

This is how the program was planned -



Over the course of the internship, my learnings enhanced my technical skills. I gained insights into project management, problem-solving, and effective communication through collaboration with mentors and peers. Each challenge encountered served as a learning opportunity

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/LoRaWAN), Java Full Stack, Python, Front end** etc.

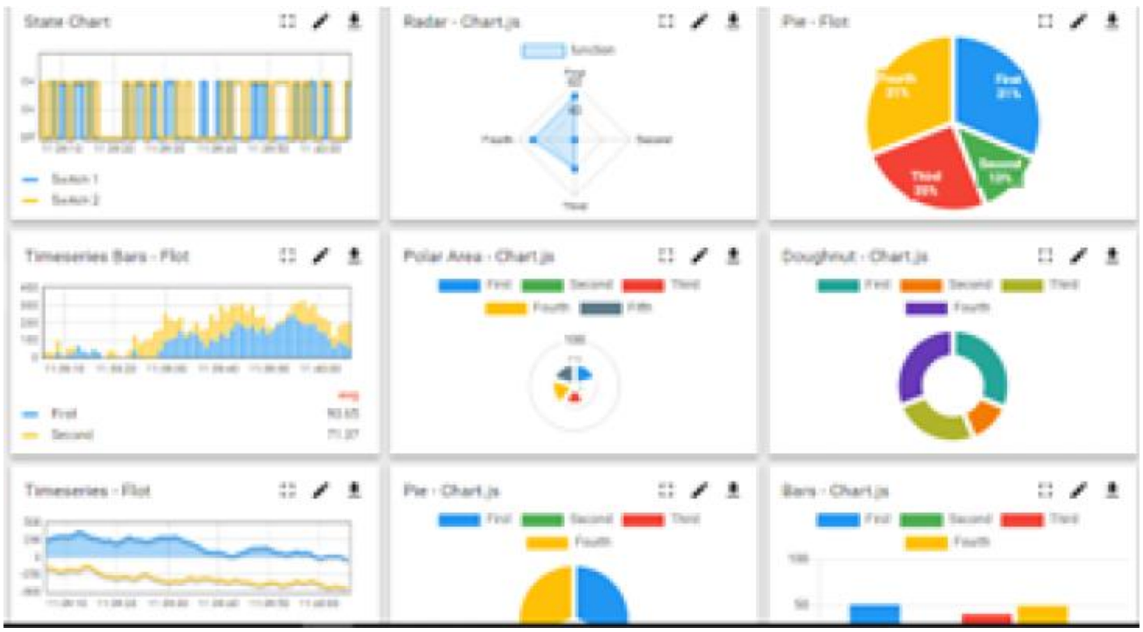
i. UCT IoT Platform (uct Insight)

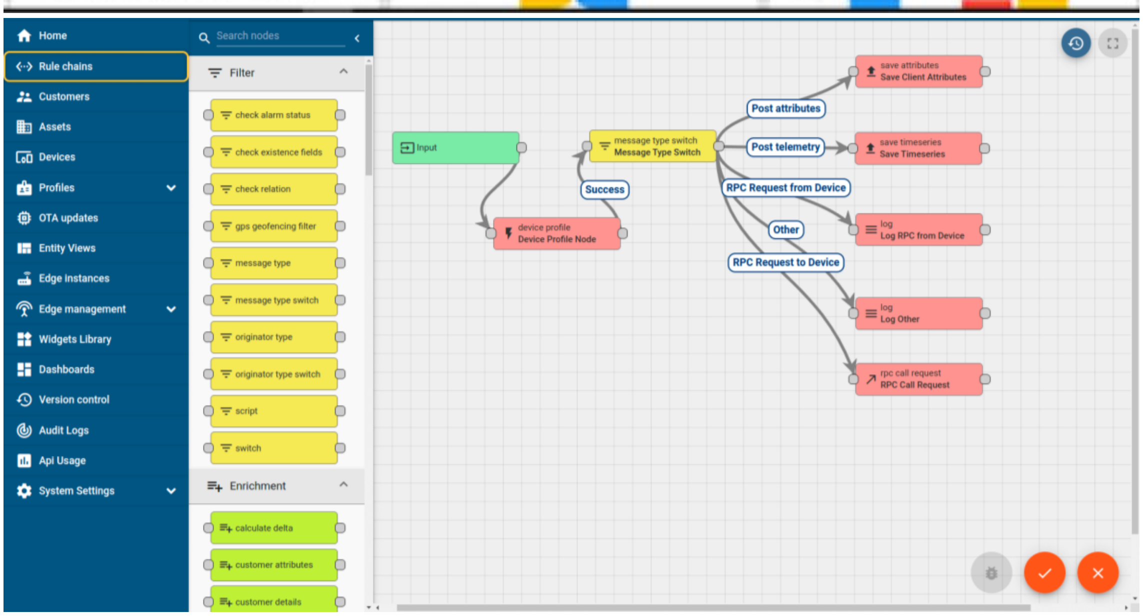
UCT Insight is an IOT platform designed for quick deployment of IOT applications on the same time providing valuable “insight” for your process/business. It has been built in Java for backend and ReactJS for Front end. It has support for MySQL and various NoSql Databases.

- 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





ii. Smart Factory Platform (FACTORY WATCH)

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 unleased 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 what 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.

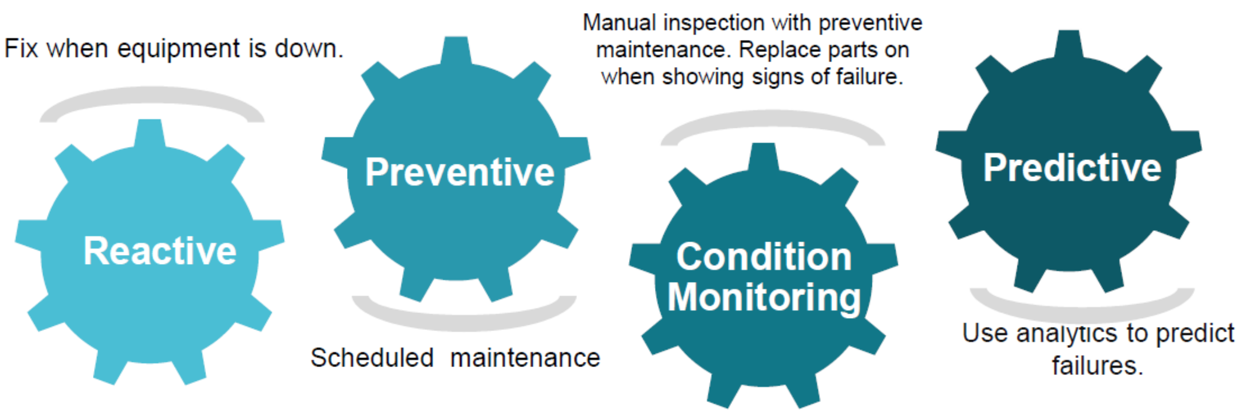
Machine	Operator	Work Order ID	Job ID	Job Performance	Job Progress		Output		Rejection	Time (mins)				Job Status	End Customer
					Start Time	End Time	Planned	Actual		Setup	Pred	Downtime	Idle		
CNC_S7_81	Operator 1	WO0405200001	4168	58%	10:30 AM		55	41	0	80	215	0	45	In Progress	i
CNC_S7_81	Operator 1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

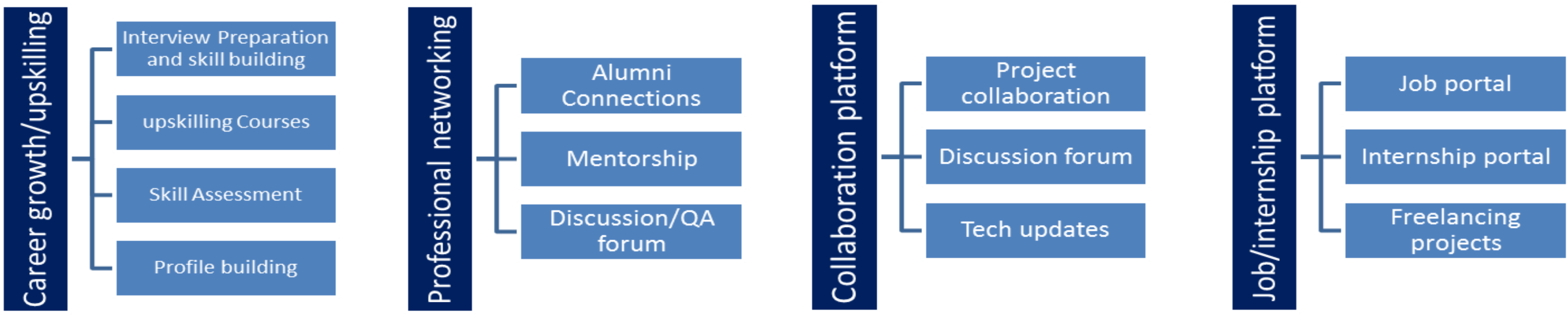
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



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.



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] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). 2016.

[2] Chollet, François. "Xception: Deep learning with depthwise separable convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). 2017.

[3] Szegedy, Christian, et al. "Going deeper with convolutions." Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). 2015. [3]

2.6 Glossary

Terms	Acronym
YOLO	You Only Look Once - A state-of-the-art object detection algorithm that applies convolutional neural networks (CNNs)
Epoch	One complete pass through the entire dataset during the training of a machine learning model.
Preprocessing	The process of preparing and cleaning data before feeding it into a machine learning model
Inference	The process of using a trained machine learning model to make predictions or classifications on new, unseen data.
Postprocessing	The processing of the output generated by a machine learning model, which may include tasks such as filtering, thresholding, and formatting the results for interpretation.
Accuracy	A performance metric used to evaluate the effectiveness of a machine learning model, representing the ratio of correctly predicted instances to the total number of instances.
F1 Score	The harmonic mean of precision and recall, providing a single metric that balances both metrics and is often used as an overall measure of a model's performance.
Confusion Matrix	A ttable that summarizes the performance of a classification model by comparing predicted class labels with true class labels, showing the counts of true positive, true negative, false positive, and false negative predictions.

3 Problem Statement

Weeds pose a significant challenge in agriculture as they compete with crops for essential resources such as nutrients, water, and sunlight, leading to reduced crop yields. Traditional methods of weed control often involve the indiscriminate use of pesticides, which can result in environmental pollution, harm to beneficial organisms, and residue accumulation in crops, posing health risks to consumers.

Problem Explanation:

Weeds are unwanted plants that grow alongside cultivated crops, competing for resources and adversely affecting crop growth and productivity. Their presence can lead to a decline in crop yields, increased production costs, and decreased farm profitability. Traditional weed control methods, such as manual labor and chemical herbicides, are often labor-intensive, time-consuming, and environmentally harmful.

Chemical herbicides, while effective in controlling weeds, pose several drawbacks. They may inadvertently harm non-target organisms, contaminate soil and water, and leave pesticide residues on crops, posing risks to human health and the environment. Moreover, indiscriminate pesticide application contributes to pesticide resistance in weeds, rendering control efforts less effective over time.

Aim:

The aim of this project is to develop a precision weed control system that selectively targets and eliminates weeds while minimizing pesticide usage on crops. By accurately identifying and targeting weeds, the system aims to reduce the environmental impact of pesticide application, mitigate the risk of pesticide residues in crops, and optimize resource utilization in agriculture.

Dataset Link:
https://drive.google.com/file/d/1MNdDKYB0x0PEW7P71bE1Jx_uLllvORA0/view?usp=sharing

The dataset provided contains images of crops and weeds captured under different environmental conditions, serving as training data for developing and validating the weed detection model. This dataset will be instrumental in training machine learning models to accurately differentiate between crops and weeds, enabling the development of a precise and effective weed control system.

4 Existing and Proposed solution

Existing solutions for crop weed detection often rely on traditional methods like manual labor or chemical herbicides. While effective to some extent, these methods have limitations. Manual labor is time-consuming, labor-intensive, and subject to human error. Chemical herbicides, on the other hand, can be harmful to the environment and human health, and may lead to the development of herbicide-resistant weeds over time.

Our proposed solution involves leveraging machine learning techniques, specifically the YOLO v8 model, to automate the process of crop weed detection. By training the model with a dataset containing images of crops and weeds, we aim to develop a system capable of accurately identifying and distinguishing between the two.

Our solution offers several key value additions compared to existing methods:

- Automation: By automating the detection process, our solution saves time and reduces the need for manual labor.
- Accuracy: Machine learning models, when properly trained, can achieve high levels of accuracy, minimizing the risk of misidentification.
- Environmental Impact: Unlike chemical herbicides, our solution is environmentally friendly, as it does not involve the use of harmful chemicals.
- Scalability: Once deployed, our system can be easily scaled to accommodate larger farming operations, improving efficiency and productivity.

Overall, our proposed solution represents a significant advancement in crop weed detection technology, offering a more efficient, accurate, and sustainable approach to weed management in agriculture.

4.1 Code submission (Github link): https://github.com/anu1808/upskillcampus/blob/master/predict.py

4.2 Report submission (Github link): https://github.com/anu1808/upskillcampus/blob/master/Anushka Behere_Final Internship Report.pdf

5 Proposed Design/ Model

Data Collection:

- Diverse dataset consisting of 1300 images of crops and weeds with YOLO labels were provided..

Data Preprocessing:

- Clean and preprocess the dataset to remove noise and irrelevant information.
- Normalize the images to ensure uniformity in lighting and color distribution.

Model Selection:

- Choose the YOLO v8 model as the primary architecture for crop weed detection.
- Consider factors such as model complexity, computational efficiency, and performance metrics.

Model Training:

- Split the preprocessed dataset into training, validation, and testing sets.
- Train the YOLO v8 model using the training set with 20 epochs, optimizing parameters to minimize loss and improve accuracy.
- Validate the model performance using the validation set, adjusting parameters as needed to prevent overfitting.

Model Evaluation:

- Evaluate the trained model using the testing set to assess its accuracy and performance metrics.
- Generate visualizations such as confusion matrices, F1 curves, and R curves to analyze model performance comprehensively.

Model Deployment:

- Deploy the trained model in a production environment for real-time crop weed detection.
- Integrate the model with a user-friendly interface for easy access and usability by farmers or agricultural professionals.

Continuous Monitoring and Improvement:

- Monitor the model's performance in real-world scenarios and collect feedback from users.
- Use feedback to identify areas for improvement and refine the model iteratively.

5.1 High Level Diagram (if applicable)

Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

6 Performance Test

The main constraint identified for the crop weed detection system is speed. In a real-world agricultural setting, the system must process images quickly to provide timely feedback to farmers. Specifically, the constraints are:

- Preprocessing time: 4.0ms per image
- Inference time: 145.7ms per image
- Postprocessing time: 1.0ms per image
- Addressing Constraints in Design:

To ensure the proposed solution meets the speed constraints, several measures were taken:

- Data Preprocessing:

Before training, extensive data preprocessing was conducted to maintain uniformity and optimize image quality. This included tasks such as noise reduction, normalization of lighting and color, and removal of irrelevant information.

By preprocessing the data beforehand, we ensured that the model receives clean and standardized input, reducing the preprocessing time during inference.

- Model Optimization:

The YOLO v8 model was chosen for its balance between accuracy and speed. It is known for its efficiency in real-time object detection tasks.

During training, model parameters were optimized to minimize inference time while maintaining high accuracy.

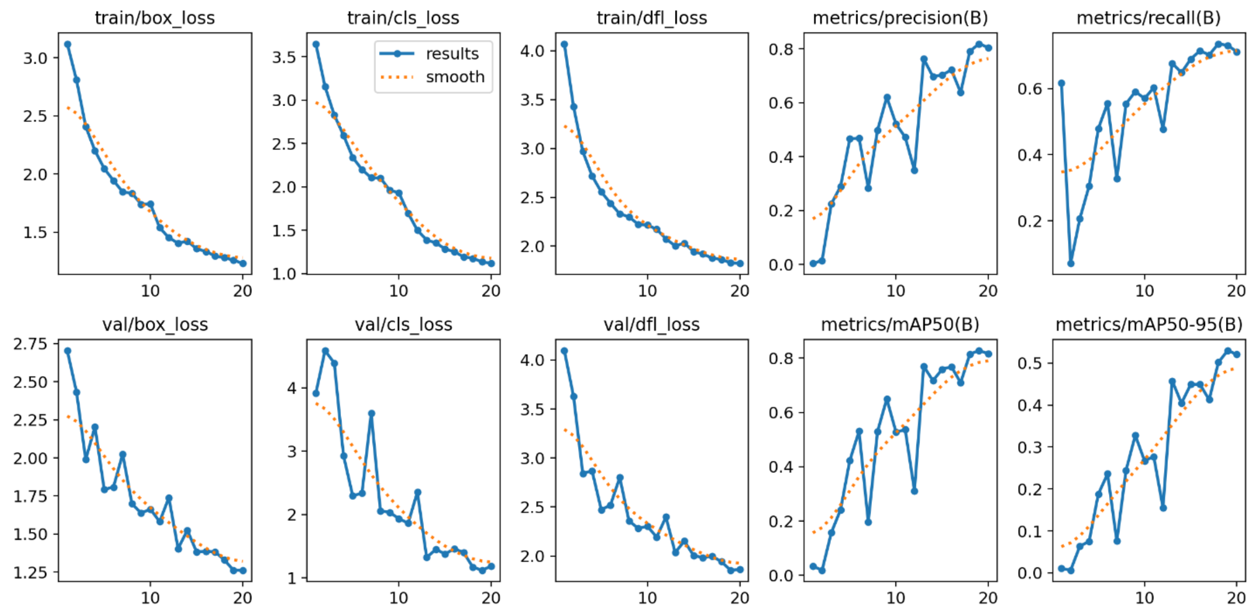
Techniques such as model quantization and pruning were explored to further reduce the model's computational load without sacrificing performance.

- Data Format Conversion:

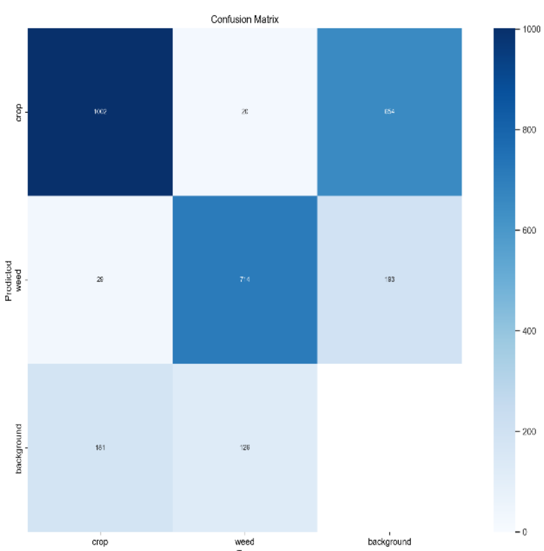
Prior to training, images were converted to the YOLO label format, ensuring compatibility with the chosen model architecture.

This streamlined the inference process, allowing the model to directly process images in the required format without additional conversions or transformations.

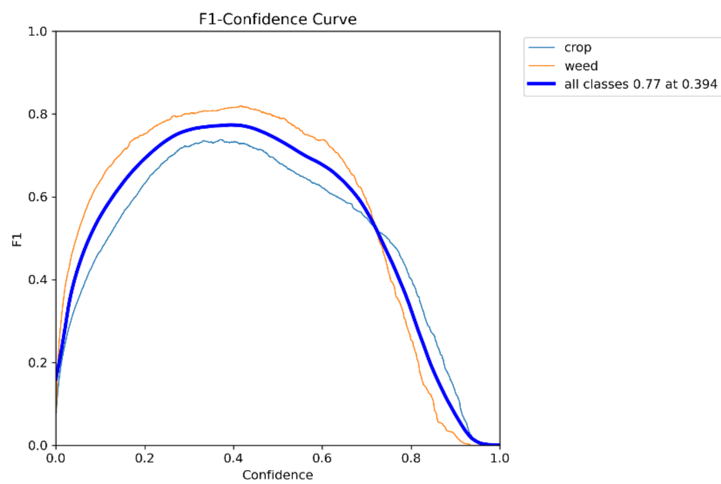
Results of Training and Validation data for 20 Epochs



Confusion Matrix



F1 Curve



6.1 Test Plan/ Test Cases

Two unseen images (1 crop and 1 weed) from the internet were selected for testing.

Test Case 1

Test Case 2

6.2 Test Procedure

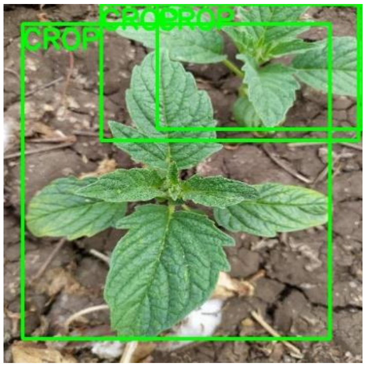
The images were fed into our model one by one, and the results were saved in the same directory.
The prediction code was written in predict.py file.

6.3 Performance Outcome

Test Case 1

Results:

- Speed:
  - Preprocessing: 8.2ms per image
  - Inference: 238.4ms per image
  - Postprocessing: 1.0ms per image
- Image Shape: (1, 3, 640, 640)



Test Case 2

Results:

- Speed:
  - Preprocessing: 4.4ms per image
  - Inference: 171.6ms per image
  - Postprocessing: 1.0ms per image
- Image Shape: (1, 3, 448, 640)

7 My learnings

Over the course of this project and internship, I've gained valuable insights and experiences that have contributed to my personal and professional growth:

- Technical Skills: Through hands-on experience with machine learning techniques, data preprocessing, and model evaluation, I've strengthened my technical proficiency in artificial intelligence and data science.

- Problem-solving Abilities: Confronting challenges throughout the project has honed my problem-solving skills. I've learned to approach complex problems systematically, breaking them down into manageable tasks and finding innovative solutions.
- Project Management: Managing the various stages of the project, from data collection to model deployment, has provided me with invaluable project management skills. I've learned to prioritize tasks, allocate resources effectively, and meet deadlines in a dynamic environment.

Career Growth:

The learnings from this internship will significantly contribute to my career growth in several ways:

- Enhanced Employability: The practical experience gained through this project will enhance my employability in the competitive field of data science and machine learning. Employers value hands-on experience and the ability to apply theoretical knowledge to real-world problems.
- Specialized Expertise: By focusing on crop weed detection, I've developed specialized expertise in agricultural technology. This niche knowledge will make me a valuable asset to organizations operating in the agriculture sector.
- Continuous Learning: The internship has instilled in me a passion for lifelong learning. I'm committed to staying updated with the latest advancements in data science and machine learning, ensuring continuous growth and adaptability in my career.

Overall, the experiences and learnings from this internship have equipped me with the skills, knowledge, and confidence to pursue a successful career in data science and make meaningful contributions to the field.

8 Future work scope

While the current project has laid a solid foundation for crop weed detection using machine learning, there are several areas for future exploration and enhancement:

Deployment on Mobile Devices:

- Explore the feasibility of deploying the crop weed detection model on mobile devices to enable real-time, on-the-go identification for farmers.
- Develop a mobile application with a user-friendly interface for easy access and usability.

Integration with Agricultural Machinery:

- Investigate the integration of the detection system with agricultural machinery such as drones or tractors equipped with cameras.
- Enable automated weed control mechanisms based on real-time detection results to optimize farming practices.

Advanced Model Optimization:

- Continuously optimize the model architecture and parameters to improve accuracy, efficiency, and scalability.
- Explore advanced techniques such as transfer learning, ensemble methods, and model compression to enhance performance while reducing computational resources.

By pursuing these future work scopes, the crop weed detection system can evolve into a powerful tool for sustainable agriculture, promoting efficient resource management, reducing chemical inputs, and increasing crop yields.