Industrial Internship Report on

”Crop And Weed Detection”

Prepared by

[Sakshi Jadhav]

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 and weed detection using YOLOV3 on agricultural image data. We know that Weed is an unwanted thing in agriculture. Weed use the nutrients, water ,land and many more things that might have gone to crops. Which result less production of required crop. Farmer often use pesticides to remove weed which also affective but some pesticide may stick with crop and may cause problem for humans. So we decided to solve these problem by developing a model which will identify which is crop and which is weed. For this, we used the dataset uploaded on kaggle.This dataset contains 1300 images of sesame crops and different types of weeds with each image labels. Each image is a 512 X 512 color image. Labels for images are in YOLO format. So this was all about our project. 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

Summary of the whole 6 weeks' work.

In this duration of internship I came across lots of problems related to my project whether it is about software , errors in code , also problems like how to increase the accuracy of the model ?, how to increase the performance of the model? . And after doing lots of research and study I found a solution over it. I also learnt lots of things related to machine learning like different algorithms in machine learning which helped to increase the accuracy & performance of the model. So this was all about the summary of the whole 6 weeks' work.

About need of relevant Internship in career development.

As this was a industrial project based internship so it plays a vital role in increasing our industrial knowledge. It helps to boost our confidence that I can work in industry very well. It also plays a big role in getting placement specially during such recession period. And I am looking forward to such industrial project based internship.

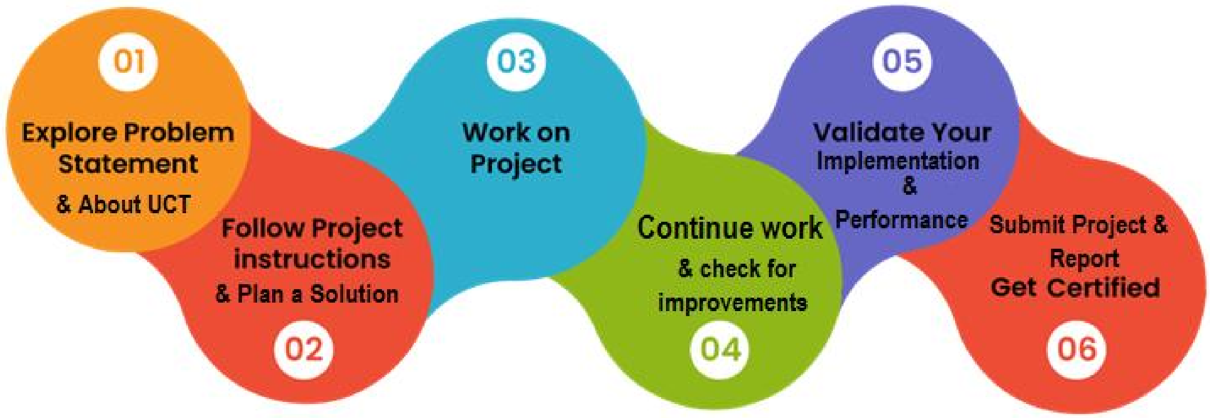
Brief about Your project/problem statement.

My project was Crop and weed detection using YOLOV3 on agricultural image data. We know that Weed is an unwanted thing in agriculture. Weed use the nutrients, water ,land and many more things that might have gone to crops. Which result less production of required crop. Farmer often use pesticides to remove weed which also affective but some pesticide may stick with crop and may cause problem for humans. So we decided to solve these problem by developing a model which will identify which is crop and which is weed.

Opportunity given by USC/UCT.

This platform provides various opportunity that helped us to widen our knowledge through quizzes, videos on various topics, etc. It gave me an opportunity to become a good engineer.

How Program was planned



Learnings and overall experience.

As I faced lots of problems during this intership like errors in code, problem while choosing the best dataset , but after doing lots of study I found out solution for it. I learnt lots of things during the duration of this internship like different algorithms in machine learning. I also learnt about the best algorithm which not only increases the accuracy of the model but also helps to increase the performance of the model.

I thank to all (Bhagyashri Saundarkar, Shruti Atkar), who have helped me directly or indirectly. Without their guidance & support it won't be possible for me to develop this project.

Message to your juniors and peers.

I will suggest all juniors & peers that this is a very good opportunity if you want to build your career in data science and machine learning. It gives us a good knowledge of the domain. As this internship is approved by the government so it will be very beneficial for placement purpose. So just go for it.

2. Introduction

a. 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.

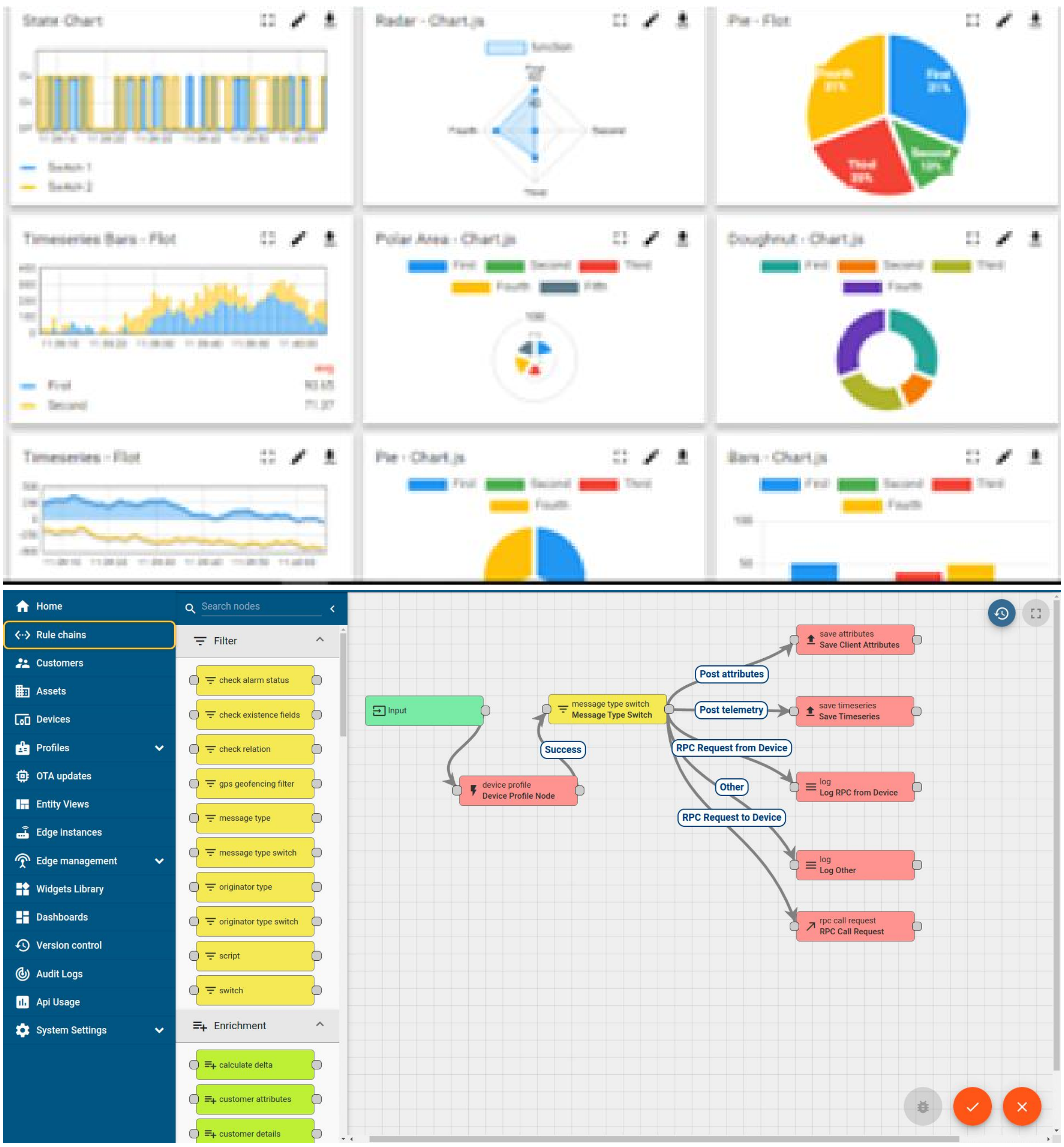
i. UCT IoT Platform (uct Insight)

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



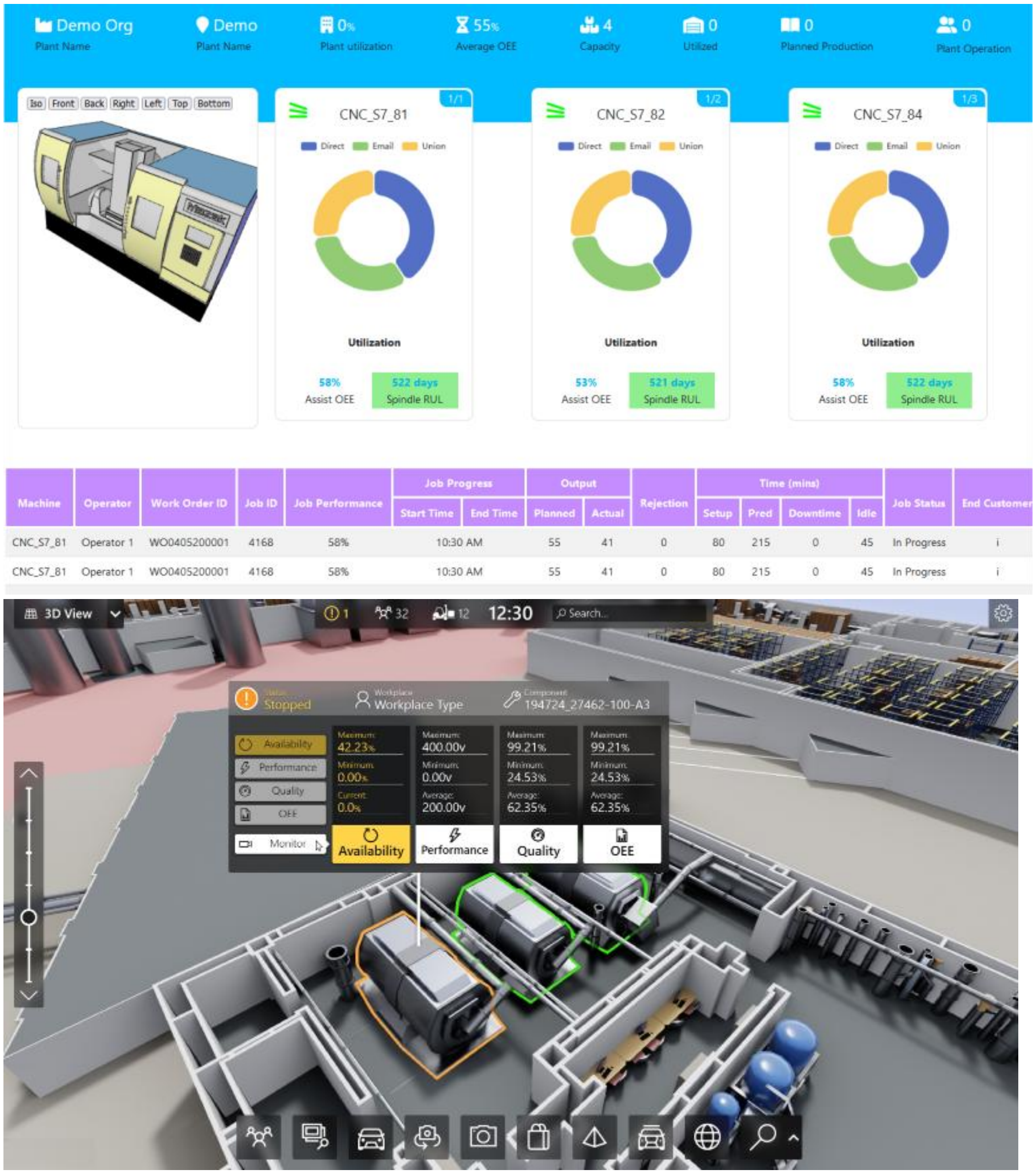
ii. Smart Factory Platform (FACTORY WATCH)

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.

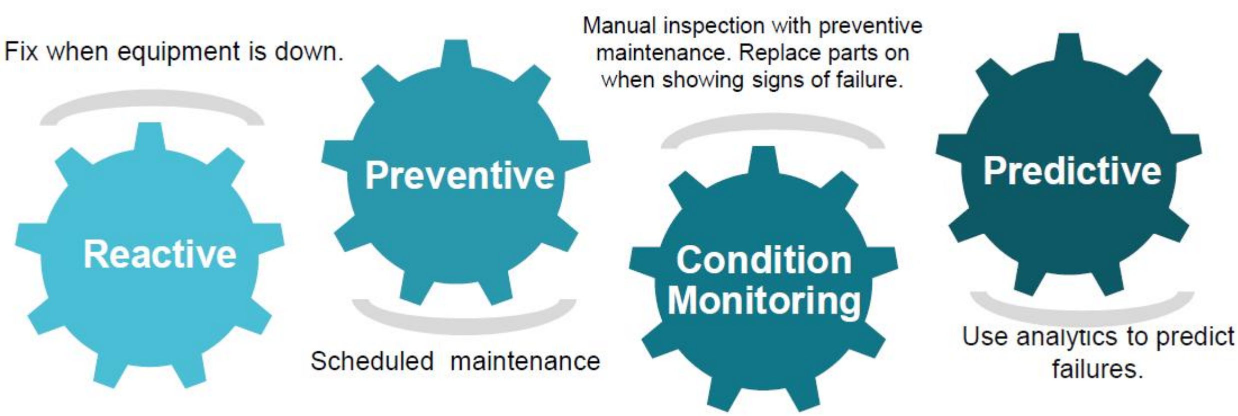

Demo Org
Plant Name
Demo
Plant Name
0%
Plant utilization
55%
Average OEE
4
Capacity
0
Utilized
0
Planned Production
0
Plant Operation
Iso Front Back Right Left Top Bottom
CNC_S7_81
CNC_S7_82
CNC_S7_84
Direct Email Union
Utilization
58%
Assist OEE
522 days
Spindle RUL
53%
521 days
Machine
Operator
Work Order ID
Job ID
Job Performance
Job Progress
Start Time
End Time
Output
Planned
Actual
Rejection
Time (mins)
Setup
Pred
Downtime
Idle
Job Status
End Customer
CNC_S7_81 Operator 1 WO0405200001 4168 58% 10:30 AM 55 41 0 80 215 0 45 In Progress
CNC_S7_81 Operator 1 WO0405200001 4168 58% 10:30 AM 55 41 0 80 215 0 45 In Progress
3D View
12:30
Search...
Stopped
Workplace Type
194724_27462-100-A3
Availability
Performance
Quality
OEE
Monitor
42.23%
0.00%
0.0%
400.00v
0.00v
200.00v
99.21%
24.53%
62.35%

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

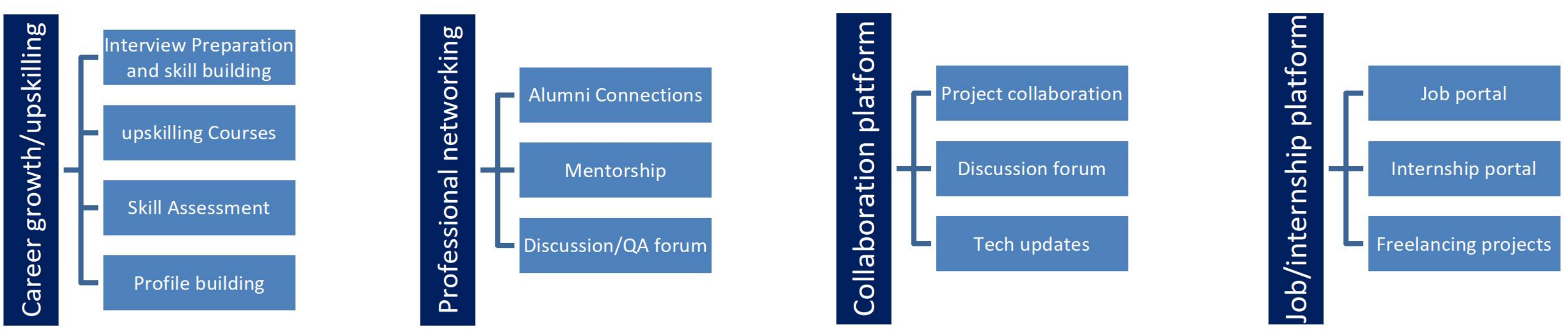
UCT is providing Industrial Machine health monitoring and Predictive maintenance solution leveraging Embedded system, Industrial IoT and Machine Learning Technologies by finding Remaining useful life time of various Machines used in production process.



b. 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.


Career growth/upskilling
Interview Preparation and skill building
upskilling Courses
Skill Assessment
Profile building
Professional networking
Alumni Connections
Mentorship
Discussion/QA forum
Collaboration platform
Project collaboration
Discussion forum
Tech updates
Job/internship platform
Job portal
Internship portal
Freelancing projects

c. 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.

d. 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.

e. Reference

For this project I took reference from various resources like IEEE research papers, youtube videos ,etc.

[1]https://deliverypdf.ssrn.com/delivery.php?ID=817078078116002007103105075091093068001008007085063041068118118105026025066000122072026035111101049057019021013027097117004116012022021021058030081100015105089070015028023095001003115107007089105094099115097092002024127083069092113107066120022118111084&EXT=pdf&INDEX=TRUE

[2] Potena, C., Nardi, D., Pretto, A. (2017). Fast and Accurate Crop and Weed Identification with Summarized Train Sets for Precision Agriculture. Advances in Intelligent Systems and Computing, 105–121. doi: https://doi.org/10.1007/978-3-319-48036-7_9

[3] Fawakherji, M., Youssef, A., Bloisi, D., Pretto, A., Nardi, D. (2019). Crop and Weeds Classification for Precision Agriculture Using Context-Independent Pixel-Wise Segmentation. 2019 Third IEEE International Conference on Robotic Computing (IRC). doi: https://doi.org/10.1109/irc.2019.00029

f. Glossary

Terms	Acronym
YOLOV3	You Only Look Once, Version 3

3. Problem Statement

We know that Agriculture is a vital sector for ensuring food security and supporting the growing global population. However, the presence of weeds in crop fields can significantly reduce crop yields, leading to economic losses and increased pesticide usage. Effective weed management is crucial to optimize crop production, minimize resource wastage, and promote sustainable farming practices. To address this challenge, the project aims to develop a machine learning-based model capable of accurately detecting and distinguishing between crops and weeds in agricultural fields.

Objectives:

- To create a robust and accurate machine learning model that can identify and classify crops and weeds in images or real-time video streams.
- To enable real-time or near-real-time deployment of the model in various agricultural settings, such as fields, greenhouses, and orchards.
- To provide farmers and agricultural professionals with a user-friendly interface for accessing detection results and recommendations.
- To improve crop management practices by facilitating targeted weed control and resource allocation, ultimately increasing crop yields and reducing environmental impact.

Scope:

The project will focus on developing a machine learning model based on the YOLOv3 architecture for object detection and classification. The model will be trained on a diverse dataset containing images and annotations of crops (various types) and common weed species. It will be designed to operate in real-world agricultural conditions, accommodating variations in lighting, weather, and crop growth stages.

4. Existing and Proposed solution

Provide summary of existing solutions provided by others, what are their limitations?

Existing Solutions:

YOLOv3-based Models: Many researchers and developers have implemented YOLOv3 or YOLOv3 variants for crop and weed detection. These models leverage the strengths of YOLO's real-time object detection capabilities.

Large Datasets: Several datasets with annotated images of crops and weeds have been created and used for training these models. Commonly used datasets include "WeedMap," "Agricultural Vision Challenge," and locally curated datasets.

Transfer Learning: Researchers often employ transfer learning by fine-tuning pre-trained YOLOv3 models on their crop and weed detection tasks. This approach allows them to leverage the knowledge gained from other datasets and domains.

Real-time Detection: Many solutions focus on real-time or near-real-time crop and weed detection, making them suitable for deployment on drones, tractors, and other agricultural equipment.

Integration: Some projects integrate these models into user-friendly applications or platforms that farmers and agronomists can use for crop management decisions.

Limitations :

1. **Limited Weed Species**
2. **Generalization**
3. **Data Annotation**
4. **Performance Under Adverse Weather**
5. **Variable Growth Stages**
6. **Resource Requirements**
7. **Data Privacy**
8. **Interpretability**

What is your proposed solution?

A proposed solution for crop and weed detection using the YOLOv3 algorithm would typically include a detailed plan for developing, training, and deploying a machine learning model to address this problem. Below is a broad outline of a proposed solution for such a project:

Problem Definition: We defined the specific objectives and scope of the crop and weed detection system. Also identified the target crops and weed species, environmental conditions, and the desired outcomes, such as improved crop management and weed control.

Data Collection and Preprocessing:

- **Data Gathering:** We Collected a diverse dataset of images or videos captured in real agricultural fields. Which includes varying lighting conditions, crop growth stages, and weed types.
- **Data Annotation:** We Annotated the dataset to label individual crops and weeds in each image or frame. This involves bounding box annotations for object localization.
- **Data Augmentation:** We also enhanced the dataset by applying data augmentation techniques, such as rotation, scaling, and flipping, to increase model robustness.

Model Selection and Architecture: Choose the YOLOv3 architecture as the basis for your crop and weed detection model. Depending on the specific requirements, you may consider modifications or variants of YOLOv3 to better suit the agricultural context.

Transfer Learning:

- **Pre-trained Models:** Utilized pre-trained YOLOv3 models, often trained on a large dataset such as COCO (Common Objects in Context), as the starting point for our model.
- **Fine-tuning:** Fine-tune the pre-trained model on our annotated dataset. This process allows the model to adapt to the specific task of crop and weed detection.

Model Training:

- **Data Split:** We Divided our dataset into training, validation, and test sets to train and evaluate the model.
- **Training:** Train the YOLOv3 model using appropriate loss functions, optimization techniques (e.g., stochastic gradient descent), and hyperparameters. Monitor the training process for convergence and model performance.
- **Validation:** Also Continuously validated the model's performance on the validation set to prevent overfitting and make necessary adjustments.

What value addition are you planning?

Real-time Disease and Stress Detection: Expanding the model's capabilities to not only detect crops and weeds but also identify signs of disease, stress, or nutrient deficiencies in plants, which can further assist in precision agriculture.

Multi-Criteria Decision Support: Developing decision support systems that consider factors like crop health, weed density, and environmental conditions to provide recommendations on optimal farming practices.

Crop Health Monitoring Over Time: Creating models that track and predict changes in crop health over time, enabling early intervention and management of crop-related issues.

a. Code submission (Github link) :

https://github.com/Sakshi757/upskillcampus.git

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

https://github.com/Sakshi757/upskillcampus/blob/main/CropAndWeedDetection_SakshiJadhav_USC_UCT.pdf

5. Proposed Design/ Model

1. Data Collection:

- Image Acquisition: Gather a diverse dataset of images or videos captured in real agricultural fields. These should include various crops, weed species, lighting conditions, and growth stages.
- Geospatial Data: Optionally, collect geospatial information (GPS coordinates) for each image to track the location of the observations.

2. Data Preprocessing:

- Data Annotation: Annotate the dataset to label individual crops and weeds in each image or frame. Use bounding boxes or pixel-level segmentation masks for object localization.
- Data Augmentation: Apply data augmentation techniques (e.g., rotation, scaling, flipping) to increase the diversity of training data and improve model robustness.

3. Model Selection:

- Architecture Choice: Choose an appropriate deep learning architecture for object detection. YOLO (You Only Look Once) variants, such as YOLOv3, are commonly used for real-time object detection.
- Pre-trained Models: Start with a pre-trained model on a large-scale dataset like COCO (Common Objects in Context) for transfer learning.

4. Model Training:

- Data Split: Divide the annotated dataset into training, validation, and test sets.
- Fine-tuning: Fine-tune the pre-trained model on your annotated dataset using appropriate loss functions, optimization techniques (e.g., stochastic gradient descent), and hyperparameters.
- Regularization: Implement regularization techniques (e.g., dropout) to prevent overfitting.

5. Model Evaluation:

- Performance Metrics: Evaluate the model's performance on the test dataset using metrics such as precision, recall, F1-score, and Intersection over Union (IoU).
- Visualization: Visualize the model's detection results on sample images to assess its accuracy and localization precision.

6. Real-time Inference:

- Deployment Infrastructure: Develop an inference engine that allows the trained model to run in real-time on edge devices (e.g., drones, cameras) or in the cloud.
- Integration: Integrate the model with a user-friendly interface or application accessible to end-users, such as farmers and agronomists.

- **Privacy and Security**
- **Testing and Validation**
- **Scaling and Optimization**
- **Documentation and Training**
- **Ethical Considerations**
- **Monitoring and Maintenance**

a. High Level Diagram

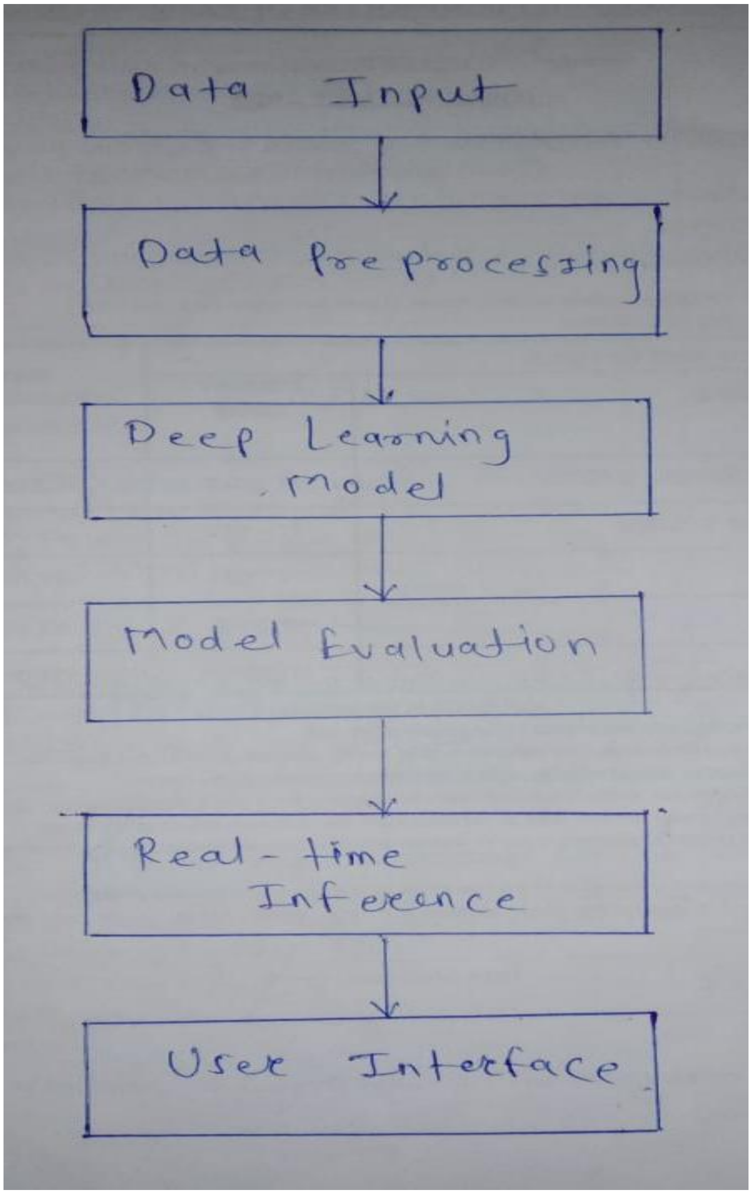
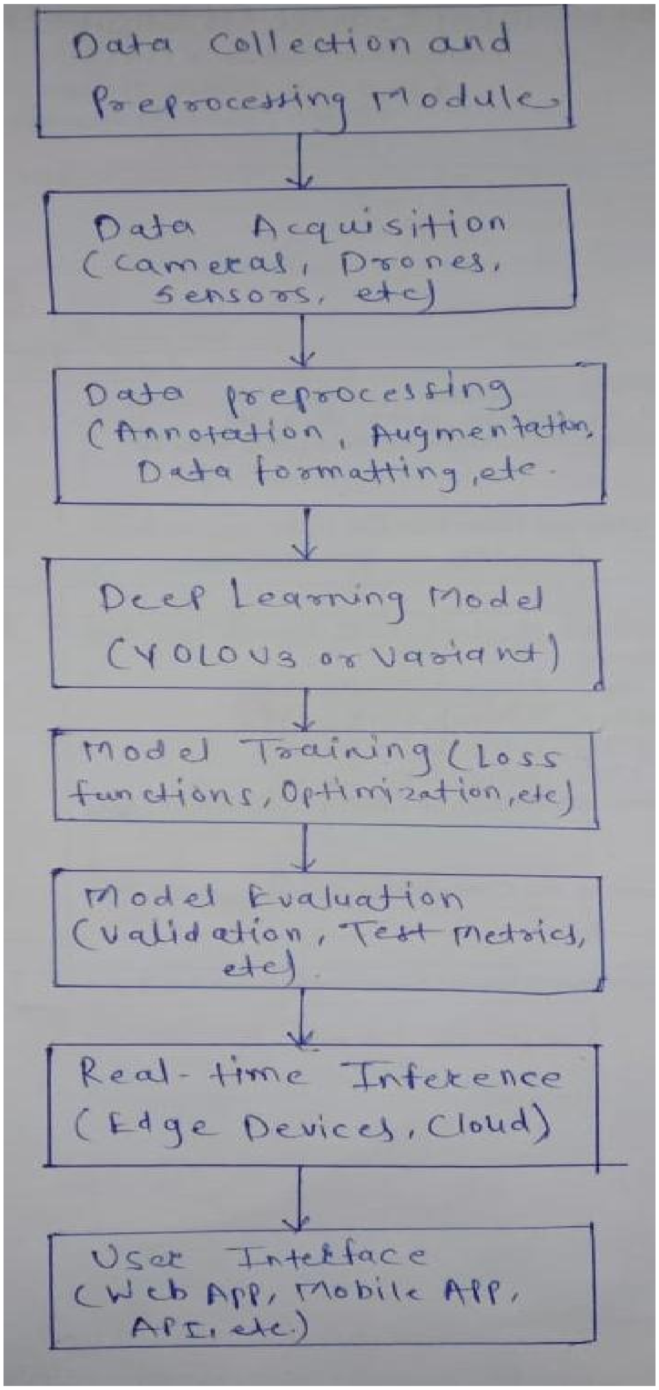


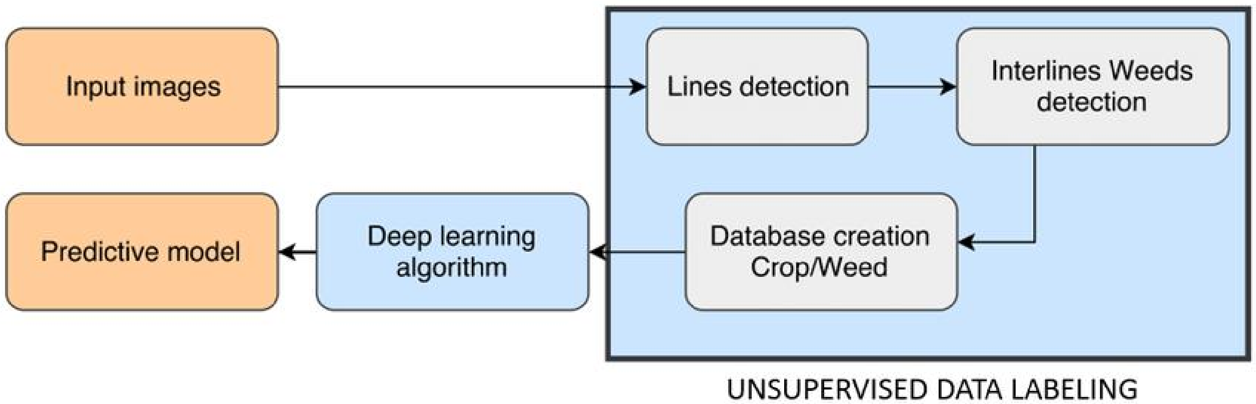
Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

b. Low Level Diagram

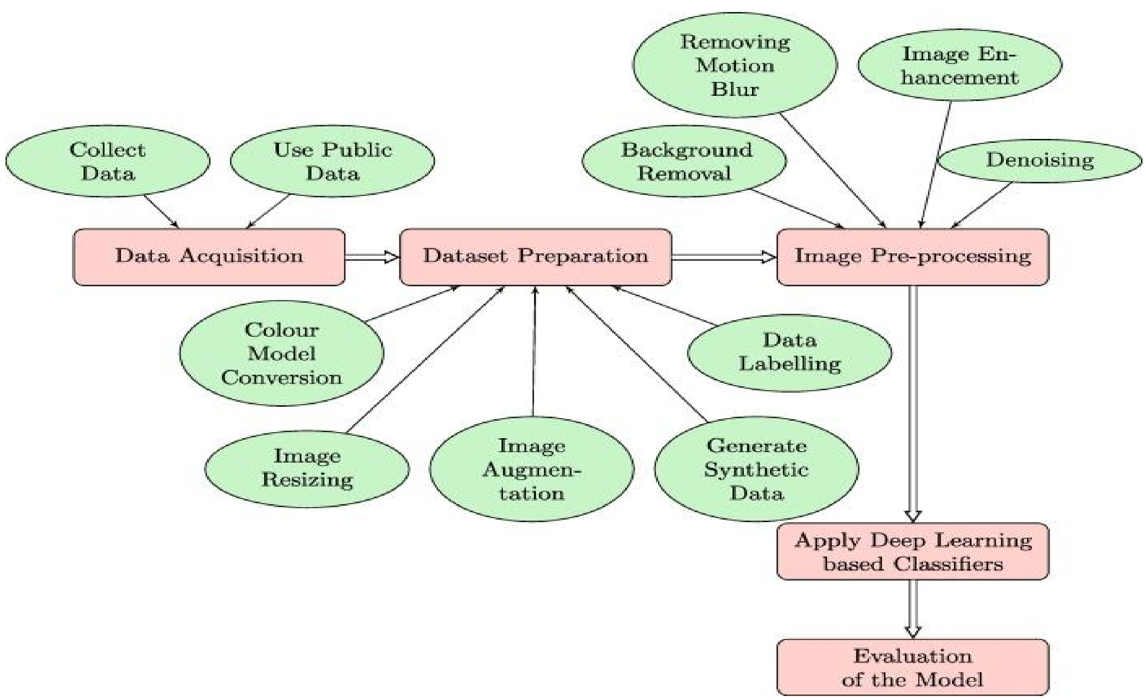


c. Interfaces

Flow Chart :



Data Flow Diagram :



6 . Performance Test

Crop and weed detection using machine learning, specifically the YOLOv3 algorithm, faces several constraints and challenges. These constraints can impact the system's performance and require careful consideration during development. Here are some of the key constraints associated with crop and weed detection using YOLOv3:

- Data Quality and Quantity
- Computational Resources
- Crop and Weed Diversity
- Data Privacy
- Scalability
- Model Interpretability
- User Acceptance
- Ethical Considerations

d. Test Plan/ Test Cases

A. Data Preprocessing Tests:

- **Data Augmentation:** Verify that data augmentation techniques (e.g., rotation, scaling) are applied correctly to increase dataset diversity.
- **Annotation Accuracy:** Ensure that annotations (bounding boxes or segmentation masks) accurately represent the location and boundaries of crops and weeds.

B.Model Training Tests:

- **Training Progress:** Monitor the training progress to ensure that the model converges and the loss decreases.

- **Overfitting Check:** Confirm that the model does not overfit by evaluating its performance on the validation set during training.

C. Model Evaluation Tests:

- **Precision and Recall:** Calculate precision, recall, and F1-score to measure the model's accuracy in detecting crops and weeds.
- **Localization Accuracy:** Evaluate the model's ability to precisely locate the center of crops and weeds.
- **Variability Testing:** Test the model's performance under different lighting conditions, weather (e.g., rain, fog), and crop growth stages.

D.Test Maintenance:

- Periodically re-run tests to ensure ongoing system performance.
- Update test cases and data as the system evolves.

e . Test Procedure

➢ Test Preparation
- Data Collection
- Model and Software Setup

➢ Data Preprocessing Tests:
- Data Augmentation Test
- Annotation Accuracy Test

➢ Model Training Tests
- Training Progress Test
- Over fitting Check

➢ Model Evaluation Tests
- Precision and Recall Test
- Localization Accuracy Test
- Variability Testing

➢ Privacy and Security Tests
➢ Data Privacy Test

- Model Security Test
- User Interface Tests
- User-Friendly Interface Test
- Feedback Mechanism Test
- Ethical Considerations
- Test Reporting
- Test Maintenance

e. Performance Outcome

The performance of a crop and weed detection system using machine learning, specifically YOLOv3, can vary based on several factors, including the quality of data, model training, environmental conditions, and the specific use case. Below are key performance aspects to consider:

- Accuracy
- Precision and Recall
- F1-Score
- Localization Accuracy
- Speed and Latency
- Robustness to Environmental Conditions
- Detection of Specific Weed Species
- False Positives and False Negatives
- Scalability
- Privacy and Security
- User Interface
- Feedback Mechanism
- Ethical Considerations

6. My learnings

You should provide summary of your overall learning and how it would help you in your career growth.

Hands-On Machine Learning: Building and training machine learning models from scratch or using existing architectures like YOLOv3 gave me practical experience in developing complex algorithms, handling data, and tuning hyperparameters.

Data Preprocessing: I had learnt the importance of data preprocessing and cleaning, which is crucial for ensuring the quality and accuracy of our model. Dealing with real-world agricultural data often involves handling noise, outliers, and missing values.

Data Collection and Labeling: Gathering and annotating datasets for crop and weed detection models can be a significant challenge. This process taught me about data acquisition strategies, data labeling, and the importance of ground-truth data.

Computer Vision: I had developed a strong understanding of computer vision techniques, including image processing, object detection, and image classification. This knowledge is highly valuable in various industries beyond agriculture.

Model Evaluation: Learning how to evaluate the performance of our models using appropriate metrics like precision, recall, F1-score, and IoU (Intersection over Union) help me gain insights into the strengths and weaknesses of our system.

Project Management: Managing a machine learning project from conception to deployment had taught me valuable project management skills, including setting goals, prioritizing tasks, and meeting deadlines.

7. Future work scope

You can put some ideas that you could not work due to time limitation but can be taken in future.

Improved Accuracy: Enhancing the accuracy of crop and weed detection models is an ongoing challenge. Future work will focus on refining model architectures, optimizing hyperparameters, and improving object detection under various environmental conditions (e.g., different lighting, weather, and field conditions).

Crop and Weed Species Recognition: Going beyond simple detection, future research can focus on classifying and recognizing specific crop and weed species. This can be valuable for tailored management strategies, as different species may require different treatments.

Fine-Grained Detection: Improving the ability to detect and distinguish between small and closely spaced plants, such as individual crop plants and weeds within rows, is an important area for advancement.

Real-Time Decision Support: Developing real-time decision support systems that not only detect crops and weeds but also suggest appropriate actions (e.g., applying specific treatments) to optimize crop yield and minimize weed competition.

Privacy and Security: Investigating methods to ensure data privacy and security, especially when using drones and other data-capturing technologies in agriculture.

Regulatory Compliance: Developing models and systems that comply with agricultural regulations and standards, especially for pesticide application and environmental protection.

Global Deployment: Adapting models for different regions, languages, and agricultural practices to promote global adoption and impact.

Education and Training: Providing training and resources for farmers and agricultural professionals to effectively use and interpret the outputs of crop and weed detection models.