Industrial Internship Report on

" Crop Weed Prediction"

Prepared by

Yashvi Baru, Nidhi Jani

| *Executive Summary* |
|---|
| This report provides details of the Industrial Internship provided by upskill Campus and The IoT Academy incollaboration 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 projectincluding the report in 6 weeks' time.

 My project was Crop Weed Prediction.

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

TABLE OF CONTENTS

1   Preface

Summary of the Whole 6 Weeks' Work - Crop Weed Prediction

Over the course of six weeks, our internship project on Crop Weed Prediction brought together a team of dedicated individuals to tackle the challenges posed by weeds in crop cultivation. Through the fusion of data science, machine learning, and agronomic expertise, we aimed to develop an innovative solution to predict and manage weeds in a sustainable and efficient manner.
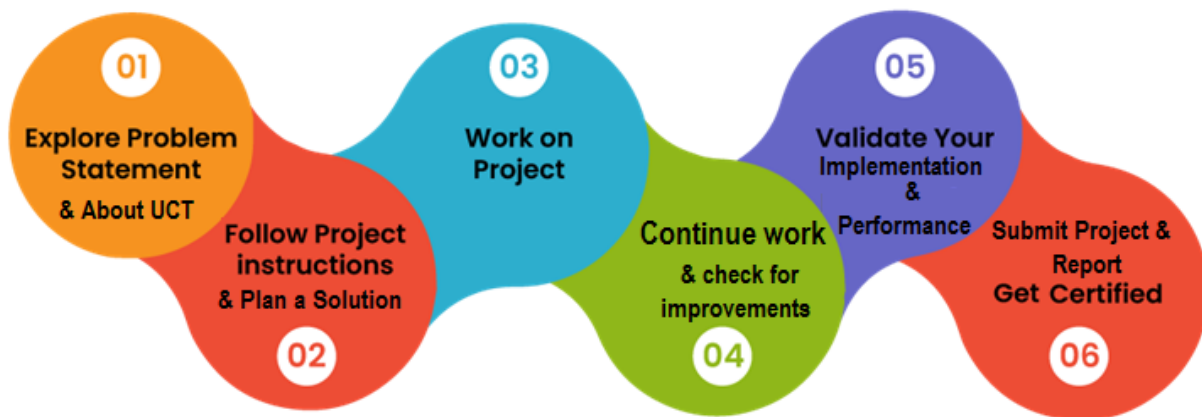
Throughout the project, we engaged in extensive research and analysis, leveraging diverse datasets such as historical agricultural data and satellite imagery. Our goal was to identify patterns and correlations between environmental factors and weed growth, enabling us to build accurate predictive models.The relevance of this internship in career development cannot be overstated. By working on such a practical and impactful project, we gained valuable hands-on experience in data analysis, machine learning, and agricultural technology. This internship provided us with a unique opportunity to apply our theoretical knowledge to real-world challenges, preparing us for future careers in the field.The problem statement for our project revolved around the detrimental impact of weeds on crop yield and food security. We recognized the need for a proactive approach to weed management, as conventional methods often rely on chemical herbicides that can have adverse effects on the environment and human health. By developing a crop weed prediction system, we aimed to provide farmers with timely and accurate information to optimize their crop management practices.This internship program, hosted by the University of Southern California (USC) and the University of Cape Town (UCT), offered a remarkable opportunity to collaborate with experts from various fields. We were able to work alongside data

---

scientists, agronomists, and software engineers, fostering a multidisciplinary environment that encouraged diverse perspectives and innovative solutions.

The program was meticulously planned, ensuring a comprehensive learning experience for all participants. We were exposed to various aspects of the project, including data collection, preprocessing, feature engineering, and model development. Regular workshops, mentorship sessions, and team meetings provided guidance and support throughout the internship.

Our learnings and overall experience during the internship were immensely rewarding. We developed a deep understanding of the agricultural landscape and the intricate relationship between environmental factors and weed growth. Through hands-on exploration of data and the application of advanced machine learning techniques, we gained valuable insights into the potential of technology to revolutionize weed management practices.

To our juniors and peers, we encourage you to seize opportunities like this internship. Engaging in practical, real-world projects enables you to bridge the gap between theory and application, fostering personal and professional growth. Embrace the challenges, collaborate with diverse teams, and never underestimate the impact you can make on the world through your skills and expertise.

As we conclude this internship, we express our gratitude to all the mentors, supervisors, and fellow interns who contributed to our growth and success. We carry forward the knowledge and experiences gained during this program, confident in our ability to contribute positively to the field of crop weed prediction and the broader domain of sustainable agriculture.

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 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
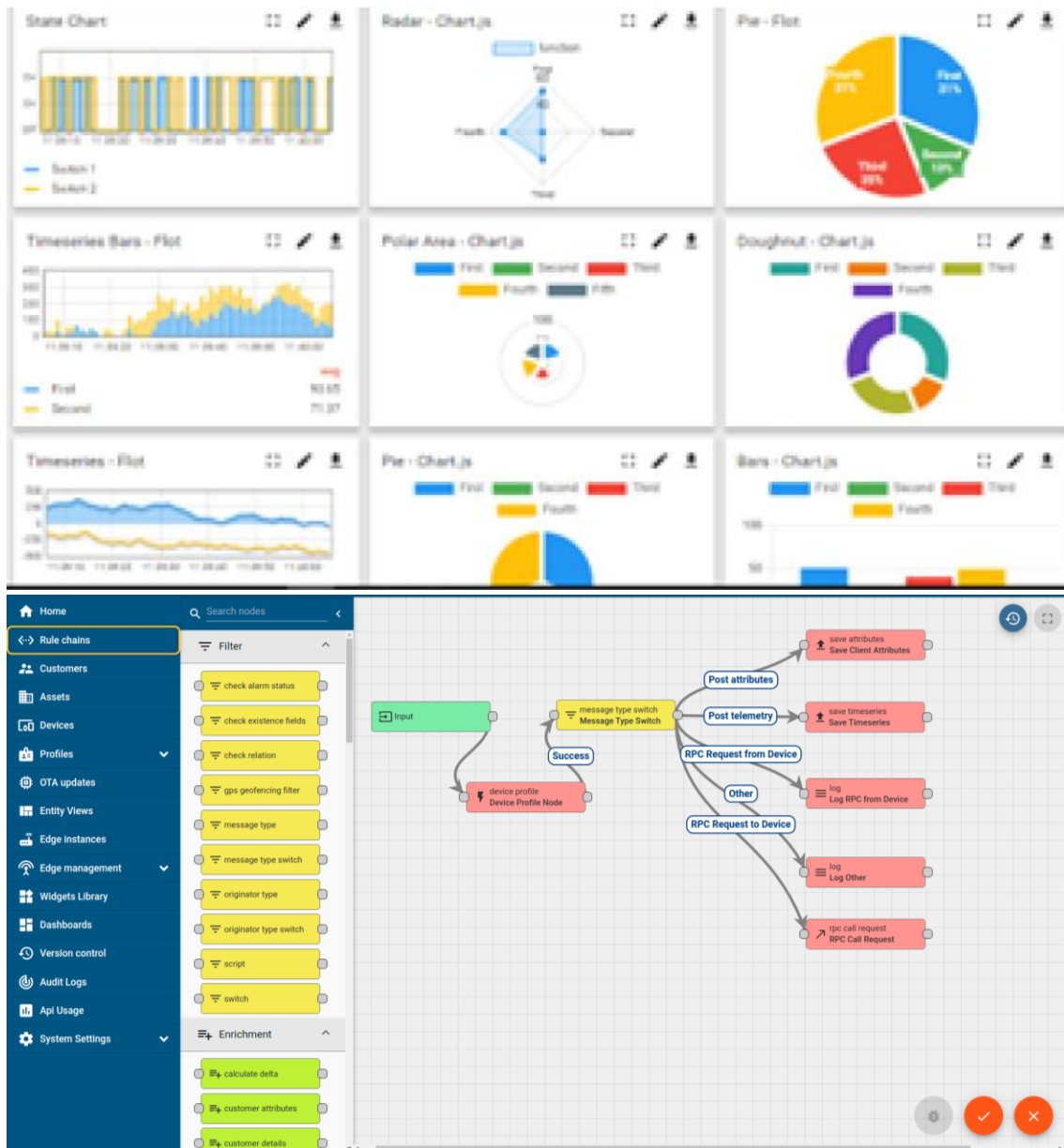
ii. Smart Factory Platform (FACTORY WATCH)

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.
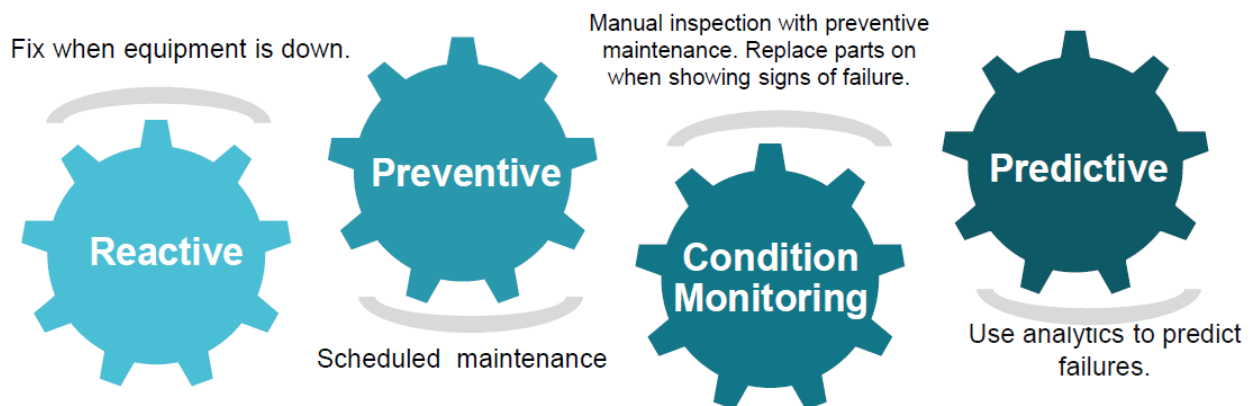
2.2 About upskill Campus (USC)

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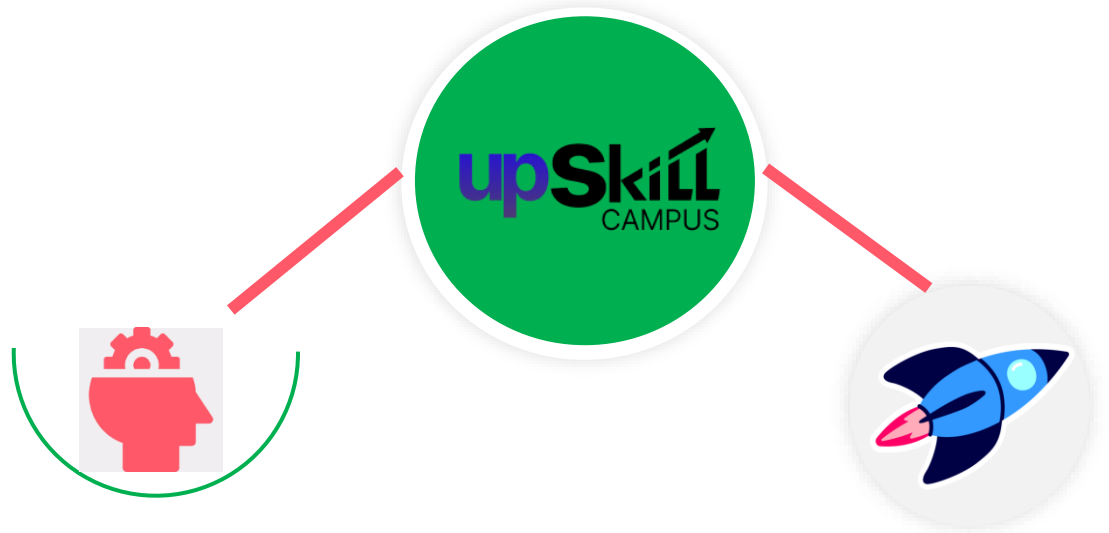
Seeing need of upskilling in self paced manner along-with additional support services e.g. Internship, projects, interaction with Industry experts, Career growth Services

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

https://www.upskillcampus.com/

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

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]   https://machinelearningmastery.com/how-to-develop-convolutional-neural-network-models-for-time-series-forecasting/

[2] https://www.researchgate.net/publication/380153793_Advancing_agriculture_with_machine_learning_a_new_frontier_in_weed_management

[3]   Using Deep Learning: A Systematic Literature Review. *Sensors* **2023**, *23*, 3670. https://doi.org/10.3390/s23073670

2.6   Glossary

Terms	Acronym
Deep Learning	DL
Machine Learning	ML
Convolution layer	Conv2D
Normalization	Norm

3 Problem Statement

In the assigned problem statement

The problem statement underscores the detrimental impact of weeds on agricultural productivity and the potential risks associated with conventional pesticide use. To address these challenges, our goal is to develop a targeted pesticide spraying system that selectively targets weeds while minimizing pesticide exposure to crops. By doing so, we aim to mitigate the adverse effects of weed competition on crop yields and reduce the risks posed by pesticide residues on harvested crops. Ultimately, our solution seeks to enhance agricultural sustainability by optimizing pesticide usage and promoting environmental and human health.

4 Existing and Proposed solution

Crop weed prediction involves using various techniques and technologies to differentiate between crops and weeds in agricultural fields, helping farmers manage weeds more effectively. Here are some of the existing solutions and their limitations:

5 Existing Solutions

1. **Manual Scouting**:

   o **Description**: Farmers or agronomists manually inspect fields to identify and map weed infestations.

   o **Limitations**:

      ▪ Time-consuming and labor-intensive.

      ▪ Subject to human error and inconsistency.

      ▪ Not feasible for large-scale farms.

2. **Remote Sensing (Satellite and Aerial Imagery)**:

   o **Description**: Utilizes satellite or drone imagery to identify weed infestations using various spectral indices (e.g., NDVI).

   o **Limitations**:

      ▪ Resolution limitations: High-resolution imagery can be costly.

      ▪ Weather dependency: Cloud cover can obstruct satellite images.

      ▪ Limited ability to distinguish between crops and weeds with similar spectral properties.

3. **Machine Vision Systems**:

   o **Description**: Uses cameras mounted on tractors or drones combined with image processing algorithms to identify and map weeds.

   o **Limitations**:

      ▪ Lighting conditions can affect image quality.

- High initial setup costs for hardware and software.

- Requires extensive training data to develop accurate models.

4. **Robotics and Autonomous Weeders**:

   o **Description**: Robots equipped with sensors and machine learning algorithms to autonomously identify and remove weeds.

   o **Limitations**:

     - High cost of robotic systems.

     - Limited by battery life and operational range.

     - Still in early stages of adoption with limited commercial availability.

5. **Machine Learning and Artificial Intelligence (AI)**:

   o **Description**: Uses deep learning algorithms trained on large datasets to identify weeds in images captured by drones or ground-based cameras.

   o **Limitations**:

     - Requires significant computational resources and expertise to develop and maintain models.

     - Data scarcity for diverse weed species and crop types can limit model accuracy.

     - Difficulty in generalizing models across different environments and crop systems.

6. **Herbicide Application Systems**:

   o **Description**: Precision sprayers that use sensors to detect weeds and apply herbicides only to affected areas.

   o **Limitations**:

     - High initial investment in precision sprayer equipment.

     - Potential for herbicide resistance development in weeds.

     - Requires accurate detection to minimize herbicide waste and crop damage.

To enhance crop weed prediction beyond the capabilities of the Random Forest algorithm, a Deep Learning (DL) model is proposed to address its limitations with the following strategies:

5. Proposed Solutions:

1. **Improved Interpretability**: Utilize techniques like feature importance analysis, model visualization, and attention mechanisms to make the complex decision-making process of DL models more transparent and understandable.

2. **Preventing Overfitting**: Incorporate regularization methods such as dropout, L1/L2 regularization, and early stopping during training to reduce model complexity and prevent overfitting, ensuring the model generalizes well to new data.

3. **Leveraging Transfer Learning**: Use pre-trained models (e.g., VGG-16) and transfer learning to take advantage of knowledge from large datasets, speeding up the training process and enhancing performance even with limited agricultural data.

4. **Handling Imbalanced Data**: Apply techniques like oversampling, undersampling, or generating synthetic samples (e.g., SMOTE) to balance class distributions, improving the model's ability to accurately predict both common and rare weed occurrences.

5. **Reducing Feature Engineering Effort**: Allow DL models to automatically learn and extract relevant features from raw data, minimizing the need for manual feature engineering and capturing complex relationships within the data.

6. **Flexible Prediction Capabilities**: Adapt DL models for both classification (e.g., weed presence) and regression tasks (e.g., weed density or crop yield) by modifying the output layer and loss function to suit the specific prediction requirements.

- **Deep Learning for Crop Weed Prediction**

Deep learning, a specialized area of machine learning, excels at training artificial neural networks to interpret complex data. In crop weed prediction, deep learning techniques, particularly convolutional neural networks (CNNs), are used to analyze images and other agricultural data to identify and classify weeds in crop fields.

Key Steps in Applying Deep Learning to Crop Weed Prediction:

1. **Data Collection**: Assemble a large and diverse dataset of labeled images showing crop plants and various weeds under different conditions and environments.

2. **Data Preprocessing**: Standardize the images by resizing and normalizing them, and use augmentation techniques like rotation and zooming to enhance the dataset's variety and robustness.

3. **Model Architecture**: Develop a CNN-based deep learning model with multiple convolutional and pooling layers to automatically learn and extract spatial features from the images, followed by fully connected layers for classification.

4. **Training**: Train the model using the labeled images. The model adjusts its parameters via backpropagation to improve its predictions by comparing them to the actual labels.

5. **Validation and Evaluation**: Validate the model on a separate set of data and evaluate its performance using metrics such as accuracy, precision, recall, and F1 score to ensure it can accurately identify and classify weeds.

6. **Deployment and Prediction**: Deploy the trained model to analyze new images of crop fields, providing predictions about the presence and types of weeds, which can be used for practical weed management.

Advantages of Deep Learning in Crop Weed Prediction:

- **Complex Pattern Recognition**: Deep learning models can discern intricate patterns and relationships within images that are crucial for distinguishing between crops and weeds.

- **High-Dimensional Data Handling**: They effectively manage and process complex and high-dimensional data typically found in agricultural scenarios.

- **Reduced Manual Feature Engineering**: Unlike traditional methods, deep learning automatically learns relevant features from raw data, minimizing the need for extensive manual input.

- **Adaptability**: These models can adapt to varying field conditions and different weed species, improving the accuracy and efficiency of weed detection and management practices in agriculture.

By leveraging deep learning, crop weed prediction models enhance the accuracy and efficiency of identifying and managing weeds, leading to more effective and sustainable agricultural practices.

4.1 Code submission (Github link)

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

5 Proposed Design/ Model

The proposed design, consisting of a well-structured CNN model, demonstrates a systematic approach to crop weed prediction. The high-level design provides an overview of the entire process, while the low-level design details the intricate architecture of the CNN. This dual-level design ensures that the model is both comprehensible and effective in achieving high accuracy in distinguishing between crops and weeds.

5.1 High-Level Diagram (if applicable)

The high-level design encompasses the overall architecture and workflow of the system, from data acquisition to prediction. The following diagram illustrates the high-level design:

1) **Data Acquisition**:

   - Collect images of crops and weeds using cameras or drones.

2) **Data Preparation**:

   - Clean and preprocess the collected images.

   - Resize images to a uniform size (512x512).

   - Augment the dataset to increase the number of training samples.

   - Label the images with bounding boxes indicating crop and weed regions.

3) **Model Training**:

   - Use the preprocessed and augmented dataset to train the CNN model.

   - Employ techniques such as batch normalization and dropout to improve model performance and prevent overfitting.

4) **Prediction and Evaluation**:

   - Use the trained model to predict on new images.

   - Evaluate the model performance using metrics like accuracy, precision, recall, and F1-score.
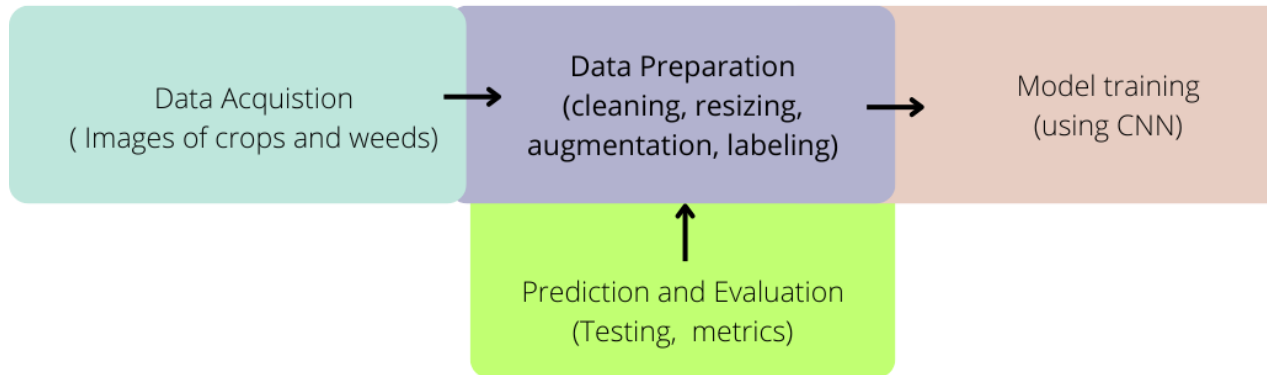
High - Level Design Diagram

Figure 1: HIGH-LEVEL DIAGRAM OF THE SYSTEM

5.2   Low-Level Diagram (if applicable)

The low-level design delves into the detailed architecture of the CNN model, including the specific layers and their configurations. The following diagram illustrates the low-level design:

1) **Input Layer**:

- Accepts input images of size 512x512x3.

2) **Convolutional Layers**:

- Multiple Conv2D layers with ReLU activation functions.

- Filters: Starting with 32 filters and increasing to 128 in deeper layers.

- Kernel Size: 3x3.

3) **Batch Normalization**:

- Normalizes the output of the convolutional layers to stabilize and accelerate the training process.

4) **Max Pooling Layers**:

- Reduces the spatial dimensions of the feature maps.

- Pool Size: 2x2.

5) **Dense Layers**:

- Fully connected layers for classification.

- First dense layer with 128 units.

- Second dense layer with 64 units.

6) **Dropout Layers**:

- Introduces regularization by randomly setting a fraction of the input units to zero during training.

7) **Output Layer**:

- A single neuron with a sigmoid activation function to output a probability score for binary classification (crop or weed).
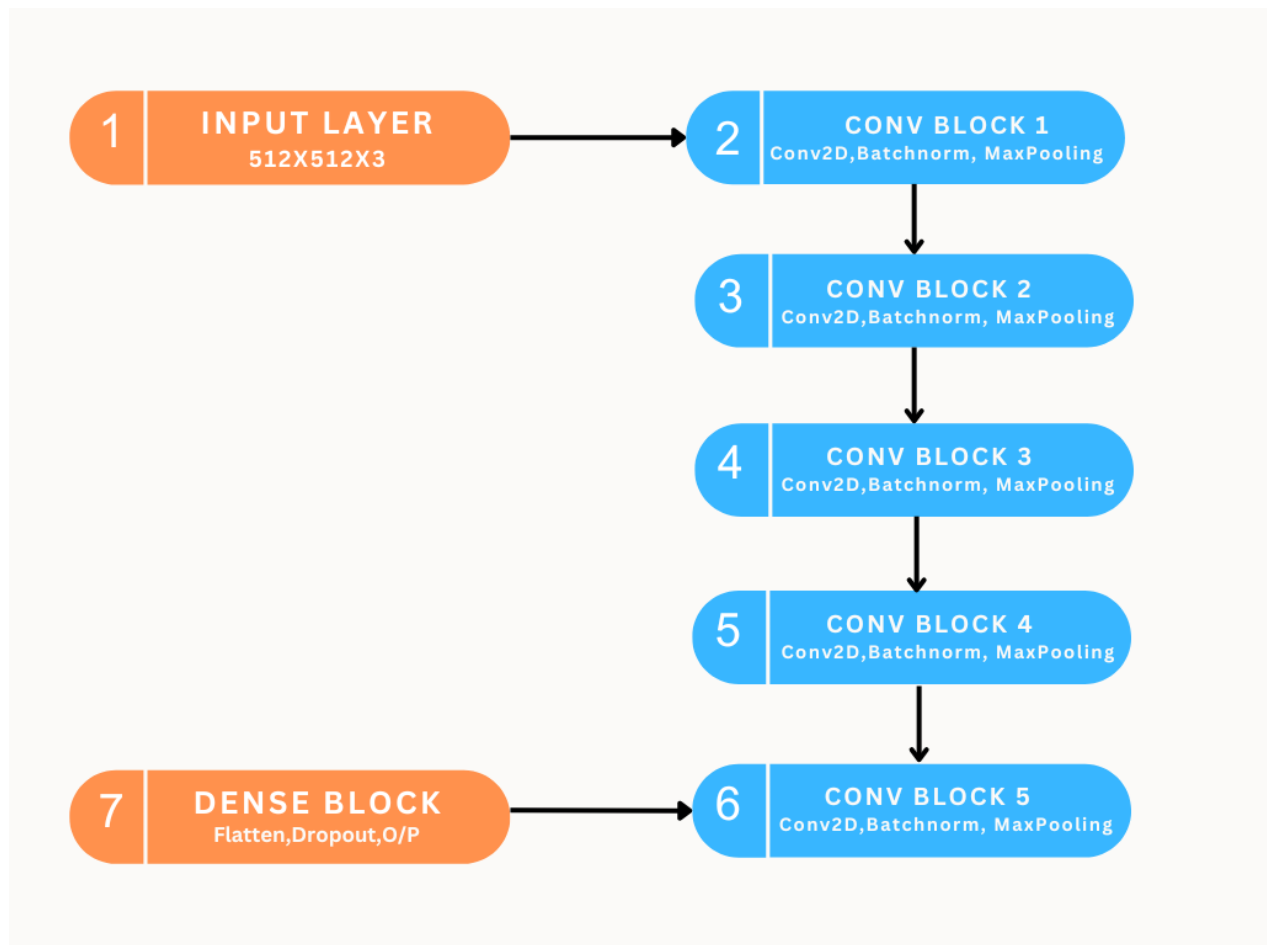
Figure 2: LOW-LEVEL DIAGRAM OF THE SYSTEM

6 Performance Test

I.   Model Compilation and Training:

To evaluate the performance of our crop weed prediction model, we compiled the model using the Adam optimizer and binary cross-entropy loss function. The model was then trained on the dataset over 10 epochs with a batch size of 32. The training accuracy was tracked to monitor the model's learning progress.

```
[81]:  # Compiling the model
       model.compile(optimizer='adam', loss='binary_crossentropy', metrics=['accuracy'])

[82]:  # Training the model
       history = model.fit(dataset, epochs=10, batch_size=32)

       Epoch 1/10
       65/65 [==============================] - 471s 7s/step - loss: 0.3799 - accuracy: 0.8431
       Epoch 2/10
       65/65 [==============================] - 462s 7s/step - loss: 0.3022 - accuracy: 0.8847
       Epoch 3/10
       65/65 [==============================] - 506s 8s/step - loss: 0.2382 - accuracy: 0.9093
       Epoch 4/10
       65/65 [==============================] - 489s 8s/step - loss: 0.2196 - accuracy: 0.9247
       Epoch 5/10
       65/65 [==============================] - 464s 7s/step - loss: 0.2090 - accuracy: 0.9276
       Epoch 6/10
       65/65 [==============================] - 443s 7s/step - loss: 0.1817 - accuracy: 0.9445
       Epoch 7/10
       65/65 [==============================] - 498s 8s/step - loss: 0.1729 - accuracy: 0.9445
       Epoch 8/10
       65/65 [==============================] - 485s 7s/step - loss: 0.1827 - accuracy: 0.9459
       Epoch 0/10
```

Figure 3

The accuracy plot below shows the training accuracy over the epochs, indicating how well the model is learning to distinguish between crops and weeds.

```
# Plotting training accuracy
plt.plot(history.history['accuracy'], color='red', label='train')
plt.legend()
plt.show()
```


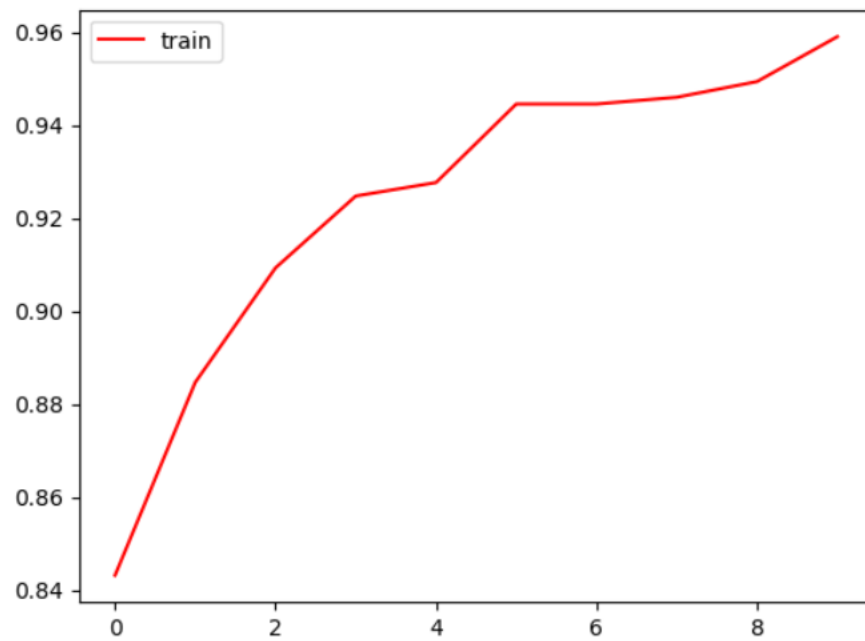**Figure 4**

6.1   Performance Outcome

To validate the performance of the trained model, we tested it with sample images from the dataset. Below is an example of testing the model with a weed image.

```
[84]:   # Testing with a weed image
        test_image = plt.imread("C:/Users/baruy/Downloads/Project5_Ag_Crop_and_weed_detection/Project5_Ag_Crop_and_weed_detection/Project5_Ag_Crop_and_weed_detec
        plt.imshow(test_image)
        plt.show()
```
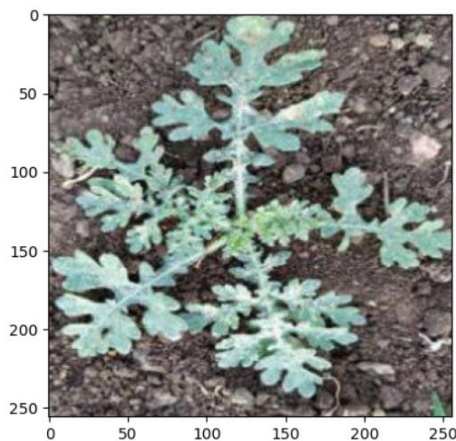


```
# Preprocessing and predicting on weed image
test_image = preprocess(test_image, 0)[0]
test_image = np.array(test_image).reshape((1, 256, 256, 3))
if model.predict([test_image])[0]:
    print("The above image is of Weed")
else:
    print("The above image is of Crop")

1/1 [==============================] - 1s 974ms/step
The above image is of Weed
```

In this step, we evaluate the model's ability to classify an image of a crop. We first load and display a crop image to verify the correct input. The image is then preprocessed to match the model's expected input format, involving resizing and normalization. The preprocessed image is reshaped to a 4D tensor and fed into the trained model for prediction. The model outputs a probability score; if this score is greater than 0.50, the image is classified as a weed. Otherwise,

it is classified as a crop. This process ensures the model's practical application in distinguishing between crops and weeds.


```
[86]:  # Testing with a crop image
       test_image = plt.imread("C:/Users/baruy/Downloads/Project5_Ag_Crop_and_weed_detection/Project5_Ag_Crop_and_weed_detection/Project5_Ag_Crop_and_weed_det
       plt.imshow(test_image)
       plt.show()

       # Preprocessing and predicting on crop image
       test_image = preprocess(test_image, 0)[0]
       test_image = np.array(test_image).reshape((1, 256, 256, 3))
       if model.predict([test_image])[0] > 0.50:
           print("The above image is of Weed")
       else:
           print("The above image is of Crop")
```

```
1/1 [==============================] - 0s 84ms/step
The above image is of Crop
```

As, in the FIG 5 we can see getting the output.

7   My learnings

Participating in the Crop Weed Prediction project has been a transformative journey, significantly enhancing my personal and professional development. Here's a detailed explanation of the key learnings and their contributions to my career growth:

1. Technical Skills Enhancement

Throughout the project, I immersed myself in a range of advanced technical tasks, including:

- **Data Preprocessing**: I learned to clean, organize, and prepare large datasets for analysis, a crucial step for any machine learning project. This involved handling missing values, normalizing data, and ensuring data quality, which are foundational skills in data science.
- **Deep Learning Model Development**: I gained hands-on experience in designing, training, and fine-tuning deep learning models, specifically tailored for image analysis in weed detection. This deepened my understanding of neural networks, particularly convolutional neural networks (CNNs), and their application in analyzing agricultural imagery.
- **Image Analysis**: I worked extensively with image data, learning techniques to process and interpret visual information to distinguish between crops and weeds. This included skills in computer vision and pattern recognition, which are vital for developing accurate prediction models.
- **Performance Optimization**: I learned to optimize model performance, balancing accuracy, and computational efficiency. This involved techniques like hyperparameter tuning, model pruning, and implementing efficient algorithms, which are essential for deploying scalable solutions.

2. Domain Expertise in Agriculture

Working on this project provided deep insights into the agricultural sector, specifically:

- **Understanding Agricultural Challenges**: I learned about the complexities of crop management, such as the impact of weeds on crop yield and the various methods farmers use to manage weeds. This knowledge is critical for developing effective, practical solutions in agricultural technology.
- **Crop and Weed Dynamics**: I gained a nuanced understanding of how different weeds affect various crops and the conditions under which they thrive. This knowledge helps in designing targeted weed control strategies that are both efficient and sustainable.

---

3. Multidisciplinary Collaboration

The project's multidisciplinary nature required collaboration with experts from various fields, which taught me several key lessons:

- **Effective Teamwork**: I worked closely with my fellow colleagues learning to appreciate and integrate diverse expertise to achieve common goals. This experience honed my ability to communicate effectively and collaborate seamlessly across disciplines.
- **Integrating Perspectives**: Understanding and incorporating different viewpoints and knowledge areas allowed us to create more comprehensive and robust solutions. This collaboration improved my ability to synthesize ideas and fostered a broader understanding of the project's objectives.
- **Project Management**: Managing timelines, coordinating efforts, and ensuring clear communication were critical aspects of the project. These experiences enhanced my project management skills, crucial for leading and participating in complex projects in the future.

4. Real-World Application and Adaptability

The project underscored the importance of applying theoretical knowledge to practical problems:

- **Translating Business Needs into Solutions**: I gained experience in understanding business requirements and translating them into actionable technical solutions. This involved identifying the needs of farmers and developing a weed prediction system that met those needs effectively.
- **Hands-On Problem Solving**: Working on a real-world problem required addressing unforeseen challenges and adapting to evolving project demands. This hands-on experience in problem-solving taught me how to navigate practical constraints and deliver workable solutions.
- **Continuous Learning**: The rapidly evolving nature of deep learning and agricultural technology emphasized the importance of staying updated with the latest tools, techniques, and methodologies. I learned to embrace continuous learning and adaptability, essential traits for staying relevant in the fast-paced tech industry.

These experiences of applying theory to practice and adapting to changing conditions will serve as a strong foundation for future projects, enabling me to deliver tangible, impactful results in my professional career.

8. Future work scope

We were unable to thoroughly investigate various ideas and approaches during the Crop Weed Prediction research due to time constraints. These concepts, however, have potential and can be followed in the future to improve the system. Here are a few examples:

Multi-Sensor Fusion: By combining data from many sensors, such as satellite imaging, drones, and IoT devices, a more comprehensive and accurate picture of the agricultural landscape can be obtained. We can improve weed prediction models and obtain better insights into the interplay of environmental factors and weed occurrences by merging data from various sources.

Long-Term Weed Growth Forecasting: Moving beyond short-term weed occurrences to long-term weed growth forecasting would allow farmers to plan and implement proactive weed management measures.

Sustainable and Efficient Farming Practices: Reduction in Chemical Usage and Support for Organic Farming: ML-based weed prediction will promote sustainable agriculture by minimizing the need for herbicides through precise application and mechanical weeding methods. This technology can also support organic farming practices by providing tools for non-chemical weed control, ultimately contributing to environmentally friendly and sustainable food production.

Integration with Advanced Precision Agriculture: Real-time Monitoring and Autonomous Systems: Future advancements will enable ML models to work in tandem with real-time data from drones, sensors, and IoT devices for precise, on-the-spot weed detection and management. Autonomous vehicles and robots equipped with these models could identify and remove weeds efficiently, reducing the need for manual labor and herbicide use.

We can develop models that predict weed growth trends and enhance long-term decision-making by taking historical data, weather patterns, and crop rotations into account.