**Industrial Internship Report on**

**“CROP AND WEED DETECTION”**

**Prepared by**

**MRIDUL KAPOOR**

|  |
| --- |
| *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 project, utilizing Convolutional Neural Networks (CNN) for effective Crop and Weed Detection.  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 4](#_Toc142994716)

[2 Introduction 5](#_Toc142994717)

[2.1 About UniConverge Technologies Pvt Ltd 5](#_Toc142994718)

[2.1.1 UCT IoT Platform 5](#_Toc142994719)

[2.1.2 Smart Factory Platform 6](#_Toc142994721)

[2.1.3 based Solution 7](#_Toc142994722)

[2.1.4 Predictive Maintenance 7](#_Toc142994723)

[2.2 About upskill Campus (USC) 7](#_Toc142994724)

[2.3 The IoT Academy 8](#_Toc142994725)

[2.4 Objectives of this Internship program 8](#_Toc142994726)

[2.5 Reference 8](#_Toc142994727)

[3 Problem Statement 9](#_Toc142994728)

[4 Existing and Proposed solution 10](#_Toc142994729)

[4.1 Biological control 10](#_Toc142994730)

[4.2 Flaming 10](#_Toc142994731)

[4.3 Chemical herbicides: 11](#_Toc142994732)

[4.4 Mechanical cultivation: 12](#_Toc142994733)

[4.5 Mulching: 12](#_Toc142994734)

[4.6 Proposed Solution 13](#_Toc142994735)

[4.7 Code submission (Github link) 13](#_Toc142994736)

[4.8 Report submission (Github link) 13](#_Toc142994737)

[5 Proposed Design/ Model 14](#_Toc142994738)

[5.1 High Level Diagram 15](#_Toc142994739)

[5.2 Low Level Diagram (if applicable) 16](#_Toc142994740)

[6 Performance Test 18](#_Toc142994741)

[6.1 Test Plan/ Test Cases 18](#_Toc142994742)

[6.2 Test Procedure 18](#_Toc142994743)

[6.3 Performance Outcome 19](#_Toc142994744)

[7 My learnings 20](#_Toc142994745)

[8 Future work scope 21](#_Toc142994746)

# Preface

Embarking on a six-week industrial internship focused on weed detection using Convolutional Neural Networks (CNN) has been an enlightening and transformative journey. This internship made possible through the collaborative efforts of upskill Campus (USC), The IoT Academy, and UniConverge Technologies Pvt Ltd (UCT), provided a platform for practical learning and real-world problem-solving. The importance of relevant internships in career development cannot be overstated. The internship served as a bridge between academia and industry, allowing me to apply my theoretical knowledge to practical challenges.

It provided a glimpse into the complexities and intricacies of working in an industrial setting, while also honing my technical skills and problem-solving abilities. The project/problem statement revolving around weed detection using CNNs presented an exciting and significant challenge. Leveraging the power of CNNs, I set out to develop a model that could accurately classify and differentiate between weed species and crops. This endeavor aimed to empower farmers with a tool that would enable them to implement selective weeding or targeted pesticide spraying, thereby minimizing pesticide use and potential health risks.

The structured program was meticulously planned, allowing me to make the most of the six weeks, from problem statement understanding to model development and evaluation. Additionally, I would like to express my appreciation to my fellow interns and peers who shared this journey with me. Their collaboration, insights, and support enriched my experience and fostered a spirit of camaraderie and collective learning.

To my juniors and peers, I urge you to embrace opportunities like this internship wholeheartedly. Seek out relevant internships that align with your interests and career aspirations. These experiences offer a unique blend of theoretical knowledge and practical application, equipping you with the skills, insights, and confidence necessary for success in the professional world.



# Introduction

## 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.



A red and blue text on a black background

Description automatically generated

### **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



### A blue text on a black background Description automatically generated

### 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 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.



### What is LoRaWAN? - Yeastar Workplace Help based Solution

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

### 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 lifetime of various Machines used in production process.



## 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.



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

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

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



## 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.

## 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.

## Reference

1. [Data Science Full Course For Beginners | Python Data Science Tutorial | Data Science With Python](https://www.youtube.com/playlist?list=PLeo1K3hjS3us_ELKYSj_Fth2tIEkdKXvV)
2. Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, Third Edition
3. Andrew Ng course

# Problem Statement

Agriculture plays a vital role in ensuring the survival and sustenance of humanity. However, it faces numerous challenges, including the proper sowing of seeds, seed selection, and protection of crops from weeds. Weeds are unwanted plants that compete with crops for resources such as nutrients, water, sunlight, and space, leading to reduced crop production. The presence of weeds poses a major challenge in agriculture, impacting the overall productivity and profitability of farms.

Weed management is crucial for farmers to ensure the successful growth of their crops. However, traditional weed control methods, such as manual removal or the use of herbicides, have their limitations. Manual removal is labor-intensive and time-consuming, often requiring significant manpower. Herbicides, although effective in weed control, can have detrimental effects on the environment, non-target plants, and potentially human health. There is a growing need for more sustainable and targeted approaches to weed management.

**Aim:**

Our aim is to design and develop a system which can detect the weeds and the crop and can easily distinguish between, then this deep learning model can be applied to various platforms like selective weeding or selective pesticide spraying which will lead to less usage of pesticides and will not stick from the crop and create health issues.

# Existing and Proposed solution

## ****Biological control****

Biologically controlling weeds depends on the simple fact that they have a natural enemy which if introduced into the system can naturally reduce and limit their growth and the size of the seed pool. The agent that is introduced can be either insects, mites, or even some specific diseases which kill them.



The biological-control measure can use bio-herbicides as the control agents, which are compounds and secondary metabolites that are obtained from microorganisms like bacteria and fungi. However, this method is not considered a long-term approach since its effects are often limited in their application duration.

Although biological control offers an environment-friendly approach to weed management, proper selection and monitoring are essential when choosing biological control which should be based on research and technical advice. This is because if the introduced agent acts as an invasive alien species, its population can explode and can even harm the agricultural crops.

## ****Flaming****

Flaming can be a highly effective weed control method, but its usage is limited to certain special circumstances since fire is a dangerous element and in the wrong hands, can do more harm than good. Flaming is famously used to control weeds growing in non-agricultural areas like roadside, railway tracks, etc.

Some of its drawbacks are accidental fires can occur due to the use of open flames, requiring strict safety protocols and monitoring. Flaming is less suitable for row crops or densely planted areas where crop plants may be at risk of heat damage. Some crops may be susceptible to burn injuries or heat damage when exposed to flames, limiting the applicability of flaming as a weed control method. Flaming can be labor-intensive and time-consuming, requiring skilled personnel and potentially increasing labor costs.



## Chemical herbicides:

Herbicides are chemical substances designed to kill or inhibit the growth of weeds. They are effective in controlling a wide range of weed species. However, drawbacks include potential environmental contamination, non-target plant damage, the development of herbicide-resistant weeds, and concerns about human health and safety during application.



## Mechanical cultivation:

This method involves physically removing weeds through plowing, tilling, or hand weeding. It is effective for larger weeds and in row crops. Drawbacks include soil disturbance, which can lead to erosion and nutrient loss, labor-intensive requirements, and the potential for re-sprouting if weed roots are not completely removed.

****

## Mulching:

Mulching involves covering the soil around plants with organic or synthetic materials to suppress weed growth. Mulch inhibits weed germination and helps retain soil moisture. Drawbacks include the cost and availability of mulching materials, potential nutrient tie-up in the soil, and the need for regular maintenance to prevent weed breakthrough.



## Proposed Solution

We have gone through all the available solutions for the weed problem in the agriculture and all have major drawbacks, like some have environmental drawbacks like health issues and insects over population in Biologically controlling weeds, some need skilled labor and are highly labor intensive, some are costly like the pesticides and flaming some are not so efficient so basically all have some major drawback.

After analyzing all the solutions, we proposed a solution which is a mixture of deep learning and mechanical engineering.

Our solution is divided into 2 parts.

* Weed and crop detection.
* A robot which can spray pesticide or uproot the detected weed in the crop.

So, our work is focused on the first part of the solution i.e. Weed and crop detection.

## Code submission (Github link)

## Report submission (Github link)

# Proposed Design/ Model

For classifying the photos as weed or crop I have used the CNN model with the following specifications of the CNN model, I have also taken the previous test weights from the image net and preprocessed the images.

Preprocessing includes.

* Removing any irrelevant images which is not a weed or a crop.
* Dividing the given images into 2 folders as weed and crop.
* Labelling each image as crops or weed.
* Compressing the images into a small size to ease the processing.
* Dividing the images as train and test
* Reading images and converting the data into a np array format

Then designed the CNN model which include these specifications.

* Sequential model
* 5 Dense layers
* 2 Flatten layers.
* ReLu and sigmoid Activation function
* DenseNet121 weights
* Loss is Binary Cross entropy.
* Optimizer is Adam with a learning rate of 0.0001.
* Metrics is accuracy.
* Trained for 50 epochs.

Then trained on a CNN model and got an accuracy of 96.67%

## High Level Diagram

A blue and green rectangle with black text

Description automatically generatedA diagram of a flowchart

Description automatically generated

Figure 1: HIGH LEVEL DIAGRAM OF THE SYSTEM

## Low Level Diagram (if applicable)

Model: "sequential"

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d (Conv2D) (None, 254, 254, 32) 896

conv2d\_1 (Conv2D) (None, 252, 252, 32) 9248

batch\_normalization (BatchN (None, 252, 252, 32) 128

ormalization)

max\_pooling2d (MaxPooling2D (None, 126, 126, 32) 0

)

conv2d\_2 (Conv2D) (None, 124, 124, 64) 18496

conv2d\_3 (Conv2D) (None, 122, 122, 64) 36928

batch\_normalization\_1 (Batc (None, 122, 122, 64) 256

hNormalization)

max\_pooling2d\_1 (MaxPooling (None, 61, 61, 64) 0

2D)

conv2d\_4 (Conv2D) (None, 59, 59, 128) 73856

conv2d\_5 (Conv2D) (None, 57, 57, 128) 147584

conv2d\_6 (Conv2D) (None, 55, 55, 128) 147584

batch\_normalization\_2 (Batc (None, 55, 55, 128) 512

hNormalization)

max\_pooling2d\_2 (MaxPooling (None, 27, 27, 128) 0

2D)

conv2d\_7 (Conv2D) (None, 25, 25, 128) 147584

conv2d\_8 (Conv2D) (None, 23, 23, 128) 147584

conv2d\_9 (Conv2D) (None, 21, 21, 128) 147584

batch\_normalization\_3 (Batc (None, 21, 21, 128) 512

hNormalization)

max\_pooling2d\_3 (MaxPooling (None, 10, 10, 128) 0

2D)

conv2d\_10 (Conv2D) (None, 8, 8, 128) 147584

conv2d\_11 (Conv2D) (None, 6, 6, 128) 147584

conv2d\_12 (Conv2D) (None, 4, 4, 128) 147584

batch\_normalization\_4 (Batc (None, 4, 4, 128) 512

hNormalization)

max\_pooling2d\_4 (MaxPooling (None, 2, 2, 128) 0

2D)

flatten (Flatten) (None, 512) 0

dense (Dense) (None, 128) 65664

dropout (Dropout) (None, 128) 0

dense\_1 (Dense) (None, 64) 8256

dropout\_1 (Dropout) (None, 64) 0

dense\_2 (Dense) (None, 1) 65

=================================================================

Total params: 1,396,001

Trainable params: 1,395,041

Non-trainable params: 960

# Performance Test

The performance of a Convolutional Neural Network (CNN) model is typically evaluated using various metrics to assess its effectiveness and accuracy in solving the problem at hand. The following are commonly used performance evaluation measures for CNN models:

* **Accuracy Graph:**
  + The accuracy graph shows how the accuracy of the model improves or changes during training.
  + At the beginning of training, the accuracy may be relatively low as the model is learning and making initial predictions.
  + As training progresses, the accuracy typically increases, indicating that the model is improving its ability to correctly classify instances.
  + The graph may show fluctuations or plateaus in accuracy, reflecting the learning process and the complexity of the problem.
  + Ideally, the accuracy should gradually increase and stabilize at a high level, indicating that the model has learned to make accurate predictions.
* **Loss Graph:**
  + The loss graph shows how the loss of the model changes during training.
  + Loss represents the error between the predicted outputs and the true labels. It quantifies how well the model is fitting the training data.
  + At the beginning of training, the loss is often high as the model makes random or incorrect predictions.
  + As the training progresses, the loss decreases, indicating that the model is fitting the data better and making more accurate predictions.
  + Similar to the accuracy graph, the loss graph may exhibit fluctuations or plateaus, reflecting the learning process and the complexity of the problem.
  + The goal is to see a gradual decrease in loss over time, indicating that the model is converging and minimizing its error.

## Test Plan/ Test Cases

The model is tested on a 30/70 basis that is 30 % is for testing and the remaining 70% for the training. So, we have 1300 images. Out of these images the 390 images are randomly selected for the test images. And the remaining 1010 images are used as the training data.

## Test Procedure

We recursively trained the model for 50 epochs with a learning rate of 0.0001 And the whole history of epochs are recorded and then mapped on the Accuracy and the loss graph for both the Train and the valid data and we got to know about the hidden pattern in the training using the graphs.

## Performance Outcome

|  |
| --- |
| ACCURACY GRAPH |
|  |
| From the accuracy graph we can see that the model initially has a very low accuracy and as we continue feeding the images its accuracy gets increased and almost reaches to 97%  This means our model can predict weed and crop correctly at the accuracy of 97% |
|  |
| **LOSS GRAPH** |
| A graph of a graph of a graph  Description automatically generated |
| From the loss graph we can see that the model initially has a very high loss in both train and Val cases and as we continue feeding the images its loss gets decreased and almost reaches to 0.2502 |

# My learnings

From this internship I got to learn and experience many things which leads to the growth of my future career, from this internship and project I learned –

* The work culture of the corporate world
* About machine learning and deep learning techniques
* Got to learn about various ML preprocessing techniques, types of machine learning, what is the job of a data engineer.
* About CNN, RNN, ANN neural networks
* How we work on images and how we can use the keras, TensorFlow and other machine learning and deep learning tools

# Future work scope

The future scope of the project is:

* Integrate with a robot and make a weed grappler based on this detection model.
* This model can be used for other type of image detections.
* Will continue the 2 parts of the project i.e., mechanical agent which will detect the weed based on this model and will either spray pesticide or will uproot the weed.