Faculty of Engineering

WEED RECOGNITION IN AGRICULTURE USING MASK R-CNN

Thah Lieng Kang

Bachelor of Engineering

Electrical and Electronics Engineering with Honours

2022

UNIVERSITI MALAYSIA SARAWAK

Grade: ______________

Please tick (√)

Final Year Project Report	✓
Masters	
PhD	

DECLARATION OF ORIGINAL WORK

This declaration is made on the 15 August 2022.

Student's Declaration:

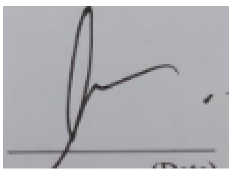
I, **Thah Lieng Kang, 67925 from the Faculty of Engineering** (PLEASE INDICATE STUDENT'S NAME, MATRIC NO., AND FACULTY) hereby declare that the work entitled **Weed Recognition in Agriculture Using Mask R-CNN** is my original work. I have not copied from any other students' work or from any other sources except where due reference or acknowledgement is made explicitly in the text, nor has any part been written for me by another person.

15 August 2022
Date submitted

Thah Lieng Kang (67925)
Name of the student (Matric No.)

Supervisor's Declaration:

I-**Tay Kai Meng** (SUPERVISOR'S NAME) hereby certifies that the work entitled **Weed Recognition in Agriculture Using Mask R-CNN** (TITLE) was prepared by the above named student, and was submitted to the "FACULTY" as a * partial/full fulfillment for the conferment of --- (PLEASE INDICATE THE DEGREE), and the aforementioned work, to the best of my knowledge, is the said student's work.



Received for examination by: ________________ Date: 15/8 2022

(Tay Kai Meng)

I declare that Project/Thesis is classified as (Please tick (√)):

☐ **CONFIDENTIAL** (Contains confidential information under the Official Secret Act 1972)*****
☐ **RESTRICTED** (Contains restricted information as specified by the organization where research was done)*****
☑ **OPEN ACCESS**

Validation of Project/Thesis

I therefore duly affirm with free consent and willingly declare that this said Project/Thesis shall be placed officially in the Centre for Academic Information Services with the abiding interest and rights as follows:

- This Project/Thesis is the sole legal property of Universiti Malaysia Sarawak (UNIMAS).
- The Centre for Academic Information Services has the lawful right to make copies for the purpose of academic and research only and not for other purpose.
- The Centre for Academic Information Services has the lawful right to digitalise the content for the Local Content Database.
- The Centre for Academic Information Services has the lawful right to make copies of the Project/Thesis for academic exchange between Higher Learning Institute.
- No dispute or any claim shall arise from the student itself neither third party on this Project/Thesis once it becomes the sole property of UNIMAS.
- This Project/Thesis or any material, data and information related to it shall not be distributed, published or disclosed to any party by the student except with UNIMAS permission.

Kang

Student signature ________________________ (15/8/2022)

Supervisor signature: ________________ (15/8/2022)

Current Address:
NO 7F, Lorong Permai Jaya 5, 96000 SIBU, SARAWAK

Notes: ***** If the Project/Thesis is **CONFIDENTIAL** or **RESTRICTED**, please attach together as annexure a letter from the organisation with the period and reasons of confidentiality and restriction.

[The instrument is duly prepared by The Centre for Academic Information Services]

WEED RECOGNITION IN AGRICULTURE USING MASK R-CNN

THAH LIENG KANG

A dissertation submitted in partial fulfilment
of the requirement for the degree of
Bachelor of Engineering
Electrical and Electronics Engineering with Honours

Faculty of Engineering

University Malaysia Sarawak

2022

ACKNOWLEDGEMENT

First and foremost, I would like to express my deepest gratitude to my supervisor, Associate Professor Dr. Tay Kai Meng for his supervision throughout the project. I am grateful to be offered this project title, weed recognition in agriculture using Mask R-CNN by my supervisor. He has been a great support to this project by giving advice and guidance to improve my project. This project could not have been completed without his assistance.

Besides, I would like to convey my appreciation to my family members for their infinite moral support throughout the project. They had also given financial support for me to complete this project without any worries on the financial aspect.

Last but not least, I am grateful to my course mates for their continuous motivations and exchange of ideas. They had been a great help throughout the project. Words cannot describe how thankful I am to those who helped me, I wish them all the best in everything they do.

ABSTRACT

Recent cooperation on deep learning has piqued the curiosity of those interested to utilise the techniques in agriculture. Weed management system is significant in agriculture that must be completed to improve crop production. The first step in weed management is to accurately classify the weeds and crops with an effective management strategy. Due to the enormous complexities in agricultural images, such as identical colour and texture, a deep neural network with pixel-wise grouping must be used to identify the plant type. The effectiveness of one of the most famous deep neural networks is examined in this paper to tackle the instance segmentation problems. Using field photos, Mask R-CNN is used to recognise weed plants (detection and classification). The dataset, which contains weeds and plants, is used to train a Mask R-CNN computer vision framework to classify and locate unique occurrences of weeds among plants. The dataset was trained on the MS COCO dataset, and the model was tailored to our classification purpose via transfer learning. Some well-reported problems in developing a suitable model are instance occlusion and the major resemblance between weeds and crops. Mask-RCNN is built on the FPN and the ResNet101 backbone. After the field images are tested on the pre-trained Mask R-CNN model, Mask R-CNN will give a class label and a bounding box offset for each weed and crop recognised. Moreover, the recognised weeds and crops will be given an object mask. Using the Mask R-CNN, the system can effectively perform instance segmentation on the images of weeds and crops with higher accuracy.

ABSTRAK

Kerjasama baru-baru ini dalam pembelajaran mendalam telah menimbulkan rasa ingin tahu mereka yang berminat untuk menggunakan teknik dalam pertanian. Sistem pengurusan rumpai adalah penting dalam pertanian yang mesti dilengkapkan untuk meningkatkan pengeluaran tanaman. Langkah pertama dalam pengurusan rumpai ialah mengklasifikasikan rumpai dan tanaman dengan tepat dengan strategi pengurusan yang berkesan. Disebabkan oleh kerumitan yang besar dalam imej pertanian, seperti warna dan tekstur yang sama, rangkaian saraf dalam dengan pengelompokan mengikut piksel mesti digunakan untuk mengenal pasti jenis tumbuhan. Keberkesanan salah satu rangkaian neural dalam yang paling terkenal dikaji dalam kertas ini untuk menangani masalah pembahagian contoh. Menggunakan foto lapangan, Mask R-CNN digunakan untuk mengenali tumbuhan rumpai (pengesanan dan pengelasan). Set data, yang mengandungi rumpai dan tumbuhan, digunakan untuk melatih rangka kerja penglihatan komputer Mask R-CNN untuk mengelaskan dan mengesan kejadian unik rumpai di kalangan tumbuhan. Set data telah dilatih pada set data MS COCO, dan model itu disesuaikan dengan tujuan pengelasan kami melalui pembelajaran pemindahan. Beberapa masalah yang dilaporkan dengan baik dalam membangunkan model yang sesuai adalah contoh oklusi dan persamaan utama antara rumpai dan tanaman. Mask-RCNN dibina pada FPN dan tulang belakang ResNet101. Selepas imej medan diuji pada model Mask R-CNN pra-latihan, Mask R-CNN akan memberikan label kelas dan kotak pembatas untuk setiap rumpai dan tanaman yang diiktiraf. Selain itu, rumpai dan tanaman yang diiktiraf akan diberikan topeng objek. Menggunakan Mask R-CNN, sistem boleh melaksanakan pembahagian contoh secara berkesan pada imej rumpai dan tanaman dengan ketepatan yang lebih tinggi.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CNN, ConvNets	-	Convolutional Neural Networks
R-CNN	-	Region-based Convolutional Neural Networks
ANN	-	Artificial Neural Networks
ROI	-	Region Of Interest
RPN	-	Region Proposal Network
AI	-	Artificial Intelligence
RGB	-	Red, Green, Blue
ReLU	-	Rectifier Linear Unit
ResNet	-	Residual Neural Network
SVM	-	Support Vector Machine
IoU	-	Intersection over Union
FC	-	Fully Connected
GPU	-	Graphics Processing Unit
YOLO	-	You Only Look Once
SSD	-	Single Shot Detector
UAV	-	Unmanned Aerial Vehicle
SLIC	-	Simple Linear Iterative Clustering
VIA	-	VGG Image Annotator
SVM	-	Support Vector Machine
HTML	-	HyperText Markup Language
COCO	-	Common Objects in Context
FPN	-	Feature Pyramid Network

ONEIROS	-	Open-Ended Neuro Electronic Intelligent Robot Operating System
API	-	Application Programming Interface
CNTK	-	Microsoft Cognitive Toolkit
CPU	-	Central Processing Unit
NumPy	-	Numerical Python
SciPy	-	Scientific Python
PIL	-	Python Image Library
VS Code	-	Visual Studio Code

CHAPTER 1

INTRODUCTION

1.1 Background

Agriculture is among the world's oldest and most important sectors. For thousands of years, society has implemented a variety of techniques, including AI, to boost agricultural performance, hence lowering negative environmental repercussions. Weed infestations cause low agricultural yields, which is one of the most important concerns that farmers face. Weed plants lower agricultural production by up to 50%, which can have serious economic consequences. Herbicides and AI-powered robots are two of the most often employed strategies for weed plant elimination. Although the former method is less expensive, there is a risk that these chemicals will harm agricultural crops, triggering a health risk. This next option, on the other hand, is costly, but it does not involve manual labour and face no health risks. As a result, the application of AI in agriculture is becoming more frequent [1].

Machine learning, deep learning, and other techniques are used by artificial intelligence to solve real-world problems. Deep learning is a type of neural network that has multiple layers of processing units and utilizes developments in processing capacity and effective training algorithms to learn complicated patterns in massive volumes of data. Additionally, Deep Learning makes use of both transformations and graph techniques to build multi-layer learning models. Deep learning algorithms developed recently have performed admirably in a variety of functions, such as audio and speech computation, visual analytics processing, and computational linguistics [2]. Deep belief networks, recurrent neural networks, and convolution neural networks are examples of deep learning architectures.

Convolution Neural Network (CNN), also known as ConvNet, has a deep feed-forward architecture and a remarkable capacity to adapt when compared to networks with fully connected layers. It has the ability to study highly abstract features and effectively detect things. First, the main reason for using CNN is the concept of weight sharing, which significantly reduces the number of parameters that must be trained, resulting in enhanced generalisation. CNN can be trained effortlessly and without overfitting because there are fewer measurements. Second, the classification step is combined with the feature extraction stage, which both utilize the learning process. Third, creating huge networks with standard ANN models is significantly more difficult than implementing in CNN. Due to their excellent performance, CNNs are widely used in a broad spectrum of applications, including image classification, object identification, face detection, voice recognition, vehicle recognition, diabetic retinopathy, and facial expression recognition [3].

Deep CNNs have been broadly applied to various object detection. CNN is a type of feed-forward neural network that works based on weight-sharing. Convolution is an interconnection that shows the way of one function coincides with the other and is a combination of two multiplied functions. The layered architecture of CNN for object detection is depicted in Figure 1.2. The image is processed with the activation function to produce feature maps. Feature maps are derived using pooling layers to reduce the geographic complexity of the algorithm. This process is iterated until the required number of filters have been applied, and feature maps have been produced as an outcome. Furthermore, these feature maps are evaluated using fully connected layers to produce image recognition output that contains a confidence score for the anticipated class labels. CNN incorporates many types of pooling layers, as indicated in table 1, to improve network complexity and minimise the number of parameters. Pooling layers are translation independent. The activation maps are sent into the pooling layers as input. They work on each patch in the chosen map[4].

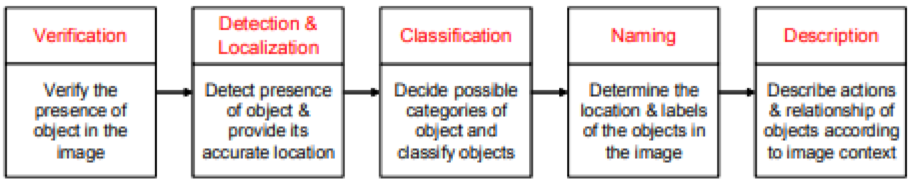


Figure 1.1: The first step in computer vision activity is object detection [4]

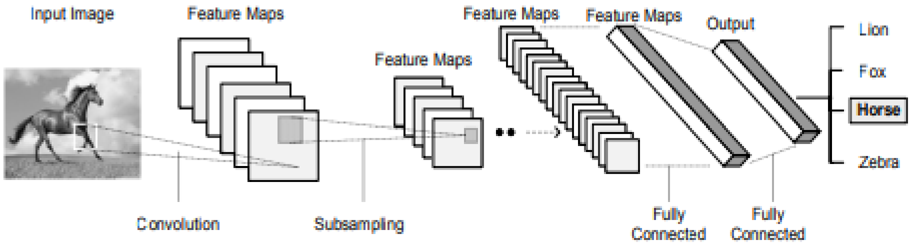


Figure 1.2: Using a CNN to Detect Objects [4]

Understanding the relationship between semantic segmentation and deep learning requires understanding that the former is not a separate study topic, but instead an evolutionary method of analysis in the progression of rough deductive reasoning to excellent deductive reasoning. Its origin dates back to computer-controlled classification methods. Categorization, in turn, can be defined as the process of anticipating an entire input, i.e., predicting the class of an item in an image or delivering a collection of classes of objects in an image based on their categorization scores. Detection, also known as object localization, is an evolutionary step from coarse to fine inference that provides not only classes of visual objects but also their location in the form of bounding boxes or centroids. The goal of semantic segmentation is to achieve fine inference by predicting labels for each image pixel. Every pixel is identified with the item or region that it is contained within. In this vein, instance segmentation assigns distinct labels to various instances of objects belonging to the same object-class. Thus, instance segmentation may be described as the job of resolving to both object identification and semantic segmentation at the same time. Part based segmentation advances this study by dissecting each segmented item into its constituent sub-components. Figure 1.3 demonstrates the progress of picture segmentation [5].

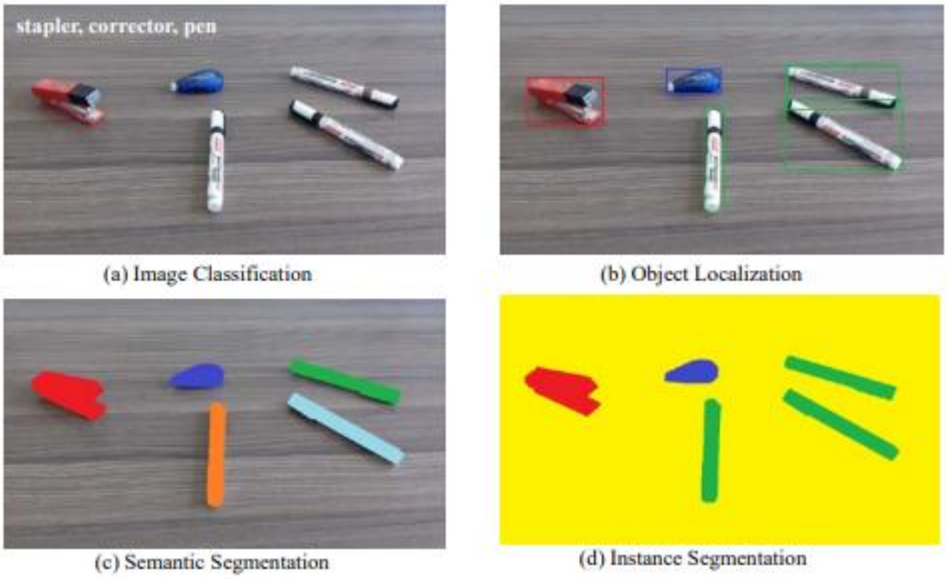


Figure 1.3: Object recognition evolution: from coarse- to fine-grained inference [5]

1.2 Problem Statement

Weed plant detection is a new research challenge in the agricultural area that seeks to use computational science to identify undesirable development of weeds in conjunction with other crops/plants.

Typically, when farmers produce anything owing to soil properties and pre-existing micro seeds, further development of weeds occurs, which detracts from the real outcome of farming by affecting the growth of planted plants. Weed detection is the challenge of properly identifying the area of weeds so that particular regions may be targeted for spraying while spraying the other plants of interest as little as possible. Precision agriculture has increasingly captured the interest of experts in recent years, as global population growth, available land, and natural resources have reduced. Approaches based on image processing might be used to overcome this challenge.

To solve the problems stated above, Mask R-CNN was employed in this work to differentiate weed plants from agricultural plants.

1.3 Objectives

The purposes of this project are:

i. To investigate the existing method of weed recognition by using the Mask R-CNN.
ii. To design and simulate the weed recognition system based on the Mask R-CNN by using the Python.
iii. To analyse the performance of weed recognition system using Mask R-CNN.

1.4 Scope

The scopes for this project are the design of the weed recognition system, implementation of the software and evaluation of the developed system. The project will mainly focus on the detection of local weeds and crop plants in the farm and field. The dataset collected will consists of images of weeds and crop plants. There are 8 classes of weeds and crop plants. In the dataset, 270 images were utilised for training and 30 images were used to validate the model. Then, the dataset trained by using transfer learning method. The project will be simulated using Python in the Visual Studio Code to show the design of the system. Design and implementation of the weed recognition system using deep learning technology with the use of Mask R-CNN, and Python. The system will be tested with the images acquired by the camera of phone. This is to test the performance of the system.

1.5 Project Outline

This project is organized into three chapters which will explained and discussed in more details. Chapter 1 covers the introduction, which are the project objectives, problem statement, summary, and main introduction of the project. The problem occurs with the weed recognition in agriculture is also included. Overall, this chapter will be focused on the brief introduction on the basic information of the deep learning and implementation of Convolution Neural Network in object detection. It also includes the thesis outline.

Chapter 2 explains on the literature review and theory during the research findings of this project. This chapter also focusing on the findings of the weed recognition system. It is also explained about the theory of CNN and R-CNN model family that will be needed in designing the weed recognition system. In this chapter also will discussed about the function of Region Proposal Network (RPN) in the architecture of the network architecture.

Chapter 3 clarifies on the methodology that will be done to finish the project successfully. The software that will be using for this project is Visual Studio Code, python and Jupyter Notebook application is presented in this chapter. In this chapter, the flowchart and the Gantt chart will be included. The theory of instance segmentation implemented on the image is also presented.

Chapter 4 present the result of projects and discussion that clearly and exactly depicts the facts and outcomes of the project. It is essentially a part of the result outcomes justification and discussion. Coding used in this project also will be discuss in this project. The problem facing by researcher will be discussed in this chapter as well as the limitation of the weed recognition system.

Chapter 5 will be the general conclusion and suggestion. In this chapter, the project research will be discussed further with recommendation for making the system more successful in the future.