ACM40960 – PROJECTS IN MATHS MODELLING 2022/23

AI Image Recognition – Weed detection

DINESH ETUKURU 20200242 SHIVAAMRUTH UPPALA 22201745

INTRODUCTION

In modern agriculture, maximizing crop yield is a pivotal challenge. The presence of **weeds** poses a significant **hindrance to crop growth**, consuming vital resources such as nutrients, water, and space that are essential for the flourishing of desired crops. While pesticides are commonly employed to combat weeds, their indiscriminate use can lead to unintended consequences, including residues on crops and potential harm to humans. In this context, our project delves into the development of an innovative solution to address this predicament. Our project centers around the creation of a comprehensive and efficient system that distinguishes between crops and weeds, enabling precise pesticide application. By leveraging **Convolutional Neural Networks (CNN)** coupled with the capabilities of **Darknet,** our aim is to develop a solution that **accurately identifies weed presence in agricultural fields.** The ultimate goal is to create a mechanism that optimizes pesticide usage by selectively targeting and eradicating weeds while minimizing any adverse impact on the crops.

METHODOLOGY – DARKNET

The methodology used here orchestrates a well-structured advancement through essential stages, aiming to create and enhance a customized object detection model. Through the seamless integration of the sophisticated **YOLO (You Only Look Once)** architecture into the **Darknet framework**, this approach is carefully designed to address the core challenge of differentiating between crops and weeds in agricultural images. Noteworthy is the incorporation of **Convolutional Neural Networks (CNN)** within the methodology, which contributes to improving the accuracy and efficiency of the detection procedure. The ultimate objective continues to revolve around enabling accurate recognition and categorization, thereby providing the foundation for effective weed management and cultivation strategies.

BLOB (BINARY LARGE OBJECT)

A BLOB IS A PRE-PROCESSED AND TRANSFORMED IMAGE THAT IS SUITABLE FOR INPUT INTO A NEURAL NETWORK FOR ANALYSIS.

FORWARD PASS

The forward pass is the process of input data flowing through a neural network to produce an output prediction.

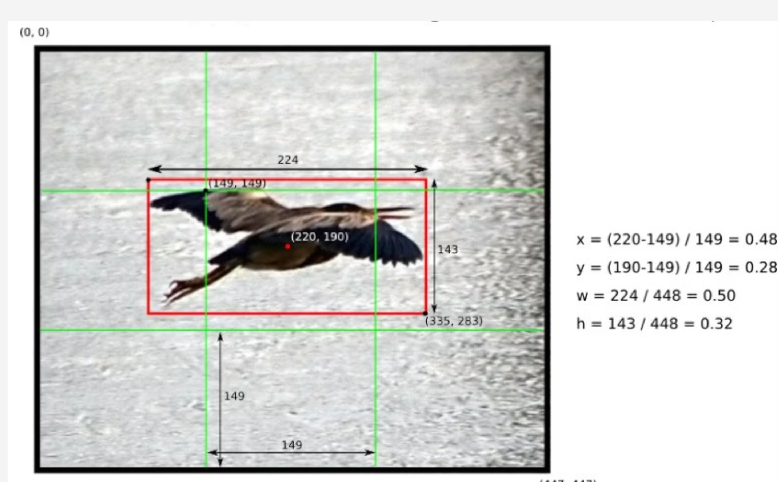
BOUNDING BOXES

Bounding boxes are rectangular regions used to define the location and extent of objects or areas of interest within an image.

OBJECT DETEDCTION IN THE IMAGE

OUTPUT FOR BOUNDING BOXES

CLASS LABEL	X	Y	Width	Height
0	0.478516	0.560547	0.847656	0.625
1	0.508789	0.489258	0.869141	0.861328



CLASS LABEL :
TWO CLASSES : 0 – CROP , 1 – WEED

X :
Normalized X-coordinate of Bounding Box's Top-left Corner

Y :
Normalized Y-coordinate of Bounding Box's Top-left Corner

Width :
Normalized Width of Bounding Box

Height :
Normalized Height of Bounding Box

HYPER PARAMETERS – TUNING :

- **1) Learning Rate :** 0.001 , 0.01
- **2) Batch Size :** 32
- **3) Number of Convolution Layers :** 53
- **4) Activation Function :** Leaky, Linear

Leaky :

$$F(X) = \max(ax, x)$$

Leaky ReLU (Rectified Linear Unit) :

We employ this approach because it is widely utilized in convolutional neural networks (CNNs) and is also prevalent in various Darknet frameworks. This choice is motivated by its ability to overcome limitations observed in traditional ReLU activation functions.

Custom Architecture Using DARKNET-53

The "Darknet-53" architecture is a deep neural network which acts as a feature extractor in YOLOv3, a prominent object detection model known for its accuracy. This deep convolutional neural network excels at extracting high-level features from images, by functioning as a crucial element within the YOLOv3 framework. Darknet-53 enables precise object detection across various scales and contexts.

	Type	Filters	Size	Output
	Convolutional	32	3 × 3	256 × 256
	Convolutional	64	3 × 3 / 2	128 × 128
1×	Convolutional	32	1 × 1	
	Convolutional	64	3 × 3	
	Residual			128 × 128
	Convolutional	128	3 × 3 / 2	64 × 64
2×	Convolutional	64	1 × 1	
	Convolutional	128	3 × 3	
	Residual			64 × 64
	Convolutional	256	3 × 3 / 2	32 × 32
8×	Convolutional	128	1 × 1	
	Convolutional	256	3 × 3	
	Residual			32 × 32
	Convolutional	512	3 × 3 / 2	16 × 16
8×	Convolutional	256	1 × 1	
	Convolutional	512	3 × 3	
	Residual			16 × 16
	Convolutional	1024	3 × 3 / 2	8 × 8
4×	Convolutional	512	1 × 1	
	Convolutional	1024	3 × 3	
	Residual			8 × 8
	Avgpool		Global	
	Connected		1000	
	Softmax			

Table 1. **Darknet-53.**

WHY CHOOSE DARKENET 53?

Feature Extraction Power:

Darknet-53 is specifically designed to excel in feature extraction from images. Its deep architecture and utilization of residual blocks make it highly effective at capturing intricate and complex patterns within images, enhancing the accuracy of object detection.

Robustness to Scale:

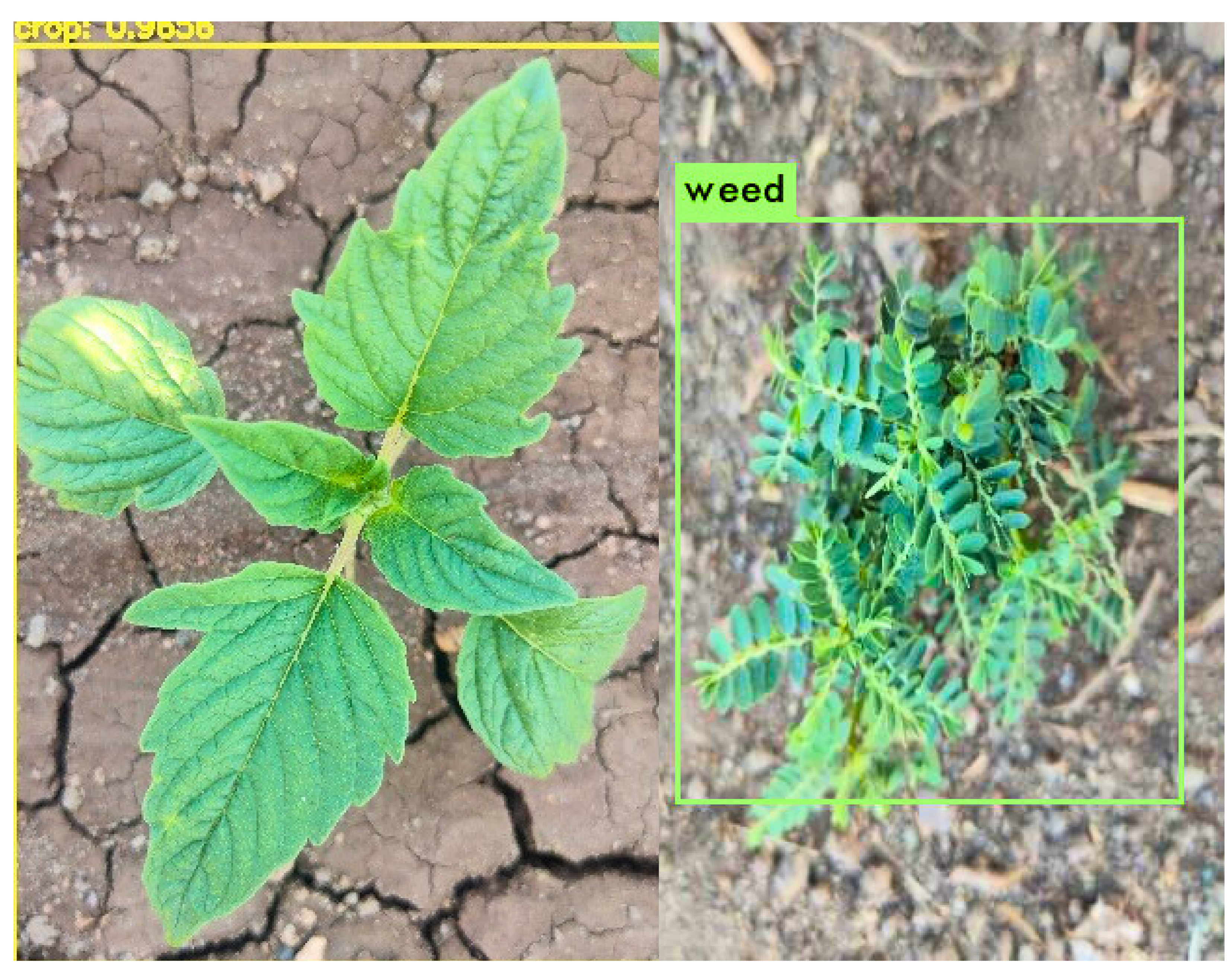
The architecture's ability to detect objects across various scales is critical in real-world scenarios where objects can appear in different sizes and contexts. Darknet-53's multi-scale feature learning capabilities enable it to handle such challenges effectively.

Performance:

Darknet-53 has demonstrated exceptional performance in accuracy and robustness when compared to other neural networks. Its architecture, inspired by successful models like ResNet, combines depth, residual blocks, and skip connections to optimize information flow and enhance detection precision.

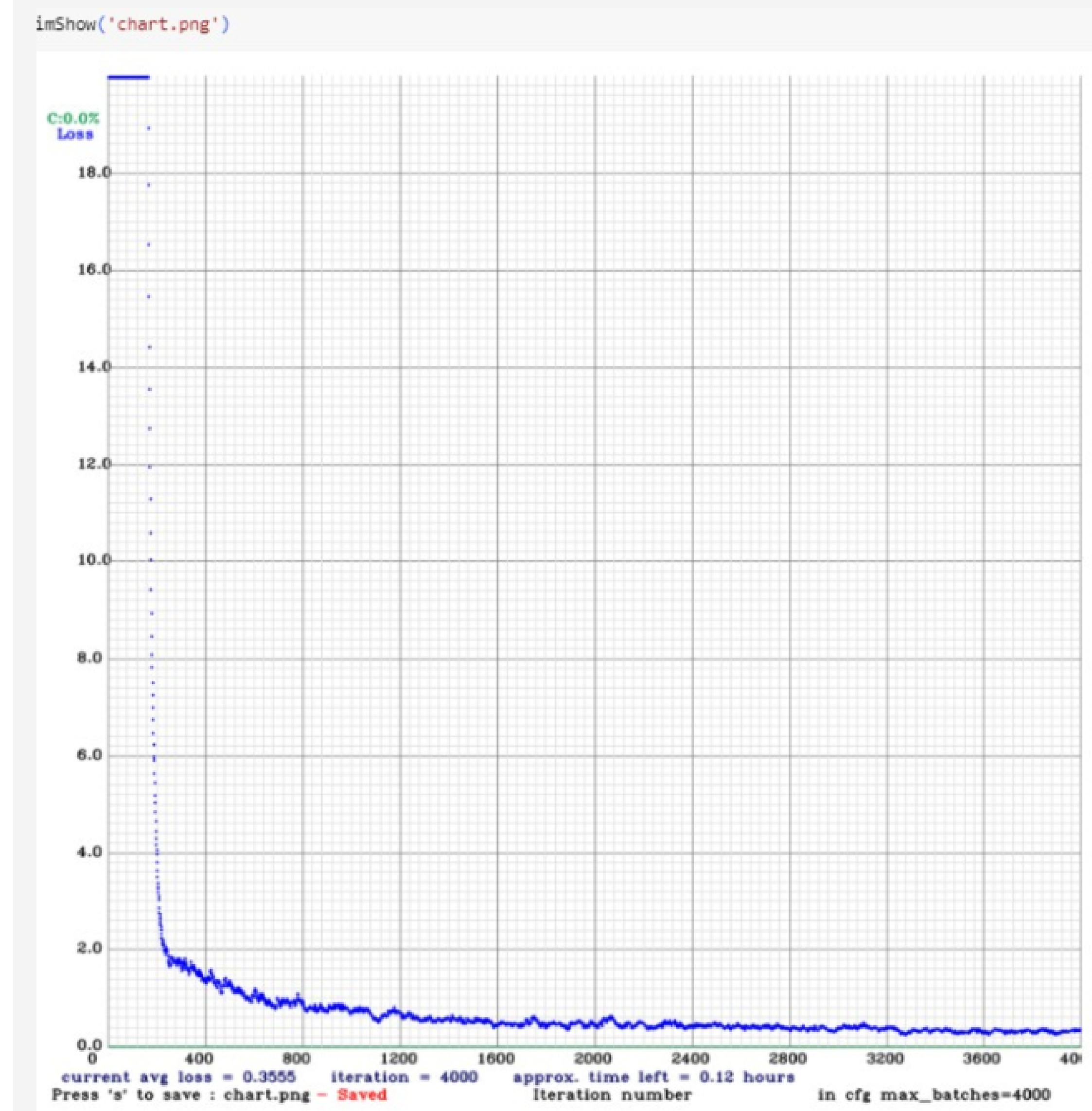
Integration with YOLOv3:

Darknet-53 is purpose-built to serve as the backbone of YOLOv3, a highly regarded and widely used object detection framework. This integration ensures seamless compatibility and optimization, leading to improved overall model performance.



RESULTS :

LOSS CURVE:



TESTING RESULTS : WEED CLASSIFICATION

The training model identifies/ detects the two objects in the image, and both of them are classified as weed with the confidence value of 0.97 and 0.38 respectively.

MAP :

MEAN AVERAGE PRECISION

CONCLUSION :

We observed a low loss curve function during the training of the model based on the Darknet-53 architecture. This indicates that the model is effectively learning from the training data. A low loss curve suggests that the model's predictions are getting closer to the actual target values (ground truth) as training progresses.

REFERENCE :

- **Redmon,J.(2015, June 8). You Only Look Once: Unified, Real-Time Object Detection. arXiv.org. https://arxiv.org/abs/1506.02640.**
- **Wang, C. Y. (2022, July 6). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv.org. https://arxiv.org/abs/2207.02696**