Development of a small scale cartesian coordinate farming robot with deep learning based weed detection

Rajalakshmi T S¹, Patrick Panikulam², Patil Kshunotra Sharad³ and Rajeev Ravindran Nair⁴

¹,²,³,⁴ Dept. of Mechatronics Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai, 603203, Tamilnadu, India

rajalakt1@srmist.edu.in

Abstract. Automated Cartesian coordinated farming is a system designed for agricultural purposes. Being one of the trends of development on automation and intelligence in the agricultural machinery, this system is able to perform certain basic elementary functions like seed sowing, spraying, watering, etc. The idea of robotics technology is being applied in agriculture. This is being designed in minimizing the labor of farmers apart from increasing the speed and accuracy of the work. A scalable Cartesian coordinate based system is modeled, which can take care of a particular area of farm land or a garden until it is time for harvesting. The system starts by planting individual seeds at predetermined locations and then automatically waters it with the exact amount required for each type of plant. It has the ability to measure soil humidity and rainfall so that water can be used depending upon on the nature of the day. The proposed system makes use of a YOLO (You Only Look Once) object detection technique to detect weeds. YOLO processes plant images at 45 frames per second in real-time, which is faster than other object detection techniques. Here, the image is divided into several grid cells before being processed. The bounding boxes as well as the class probabilities are predicted by one single neural network, in a single evaluation. This effectively boosts the speed and accuracy of weed detection.

1. Introduction

Agriculture is a highly wasteful, unpredictable expensive industry. Thus smart automated farming technology does have a lot of potential to improve the efficiency of this process. Such agriculture equipment is available in the form of huge heavy machinery. Smart agriculture can be achieved by involving advanced technology into existing farming methods. This increases production, efficiency as well as the quality of the produced product. Automated farming using a Cartesian coordinate based setup is done to reduce the difficulty of maintaining a farm for small or large scale alike by automating the basic and time consuming farming tasks such as sowing of seeds, watering of plants and removing weeds. The Cartesian coordinate farming system hardware employs linear guides or paths in the X, Y, and Z directions. This requires for tooling such as seeding nozzle, watering nozzles, humidity and soil probes, and weed removal wedges to be perfectly positioned and used on the plants and soil bed.

Different plants need different soil conditions, amount of water for growth and intervals between each watering sessions. When different kinds of plants are grown together, manually caring for each kind of plants could become cumbersome and time consuming. But a Cartesian coordinate based system knows precisely where each kind of seeds are planted and can care for each plant individually.
The interchangeable tool and sensor holder makes for a simple and convenient method for testing and preparing the soil.

The weeds that grow along with the plants must be removed at regular intervals for the proper growth of the plants. The Weed detection is done using Deep Learning Object Detection technique. The camera mounted on frame of the Z axis will be used to acquire images of the farm bed. The plant image is captured and checked whether the plant is the required crop or not. If not, the plant is considered as a weed. The position of the weed is then mapped to real world coordinates and is removed using a weeder.

2. Related works

2.1. Cartesian coordinate based automated farming

Cartesian coordinate based automated farming involves design of mechanical system. The design mechanism for automated farming as by [8] takes parameters that the plant needs for it to grow vigorously along with the automation tools which give an output requirement of fertilizing, watering etc. Besides the control system on watering, System design is also required to place plants. The system requires sensor parameters as input. Apart from the automation system there is a monitoring system along with it to monitor these parameters in real time conditions. The mechanical system consists of three axes, the X, Y and Z. The components for the system include Plant Container, Tracks and Electronic Circuits. Here, stepper motor is used for motion control in each of the axis and a microcontroller for data processing and controlling stepper motor.

Considering the rise in the demand for food, the efficiency of agricultural productivity is estimated to increase by 25% by the year 2050 [1]. Therefore, it is more important to focus on the issues that are faced by this agricultural industry. [4] proposes different categories like soil management, water management, livestock management, weed detection etc in order to classify the challenges that are faced by machine learning in the domain of precision agriculture. New developments would actually assist in the facing of the biological threat in the crop production. According to [5], it is seen that weeds create an average of 34% of production loss as the crops are made to compete for water, sunlight and minerals. Moreover, it is difficult to detect weeds, which owes to the fact of their overlapping with other crops, as well as their irregular growth position.

2.2. Object detection

Manual weeding was the oldest technique used in-order to control weeds in crops. But, actually it is more laborious as well as time consuming. This makes it in-efficient for large-scale crops. Now-a-days, agricultural industry has more of chemical weeding system, and lesser mechanic weeding system. But in this case of Andean Highlands, 75% of the vegetables say lettuce has manual weeding. This results in an inefficient production which is expensive as well[6]. Also, in the case of weeding, there is usually high margin of error. This eventually results in the damage of the plants [4]. Previous works of [2] and [7] in the lettuce field obtained good results for identification of weed as well as crop, using both the RGB as well as multispectral images. But in case of weeds in the highland tropical conditions, weeds are found to have different forms; plus, they grow in patches of bigger size. This makes the detection even more difficult.

The weeds were detected by capturing the image from the field. Then using convolutional network, the image is classified as “Crop” or “Weed” where weed indicates the presence of weed in the image. A custom architecture for the Convolutional Network was used where the input was 100*100 image and the output layer was a softmax layer that is used to classify whether the image was “Crop” or “Weed”. Also data augmentation was done on the image to improve the accuracy of the model. Once the image is classified as “Weed”, then the weed is removed by using a weeder.
3. Proposed model
In the following section, the proposed crop-weed detection model is described. The detection system consists of dataset collection, data augmentation, object detection (crop as well as weed), and the post processing stage.

3.1. Design and fabrication
The machine is designed to take care of a grow volume of 990 x 745 x 500 mm. The machine’s mechanical component consists of 2020/2040 aluminium extrusions, their supporting corner brackets and 3D printed components and fixtures. The brackets, fixtures and carriage those had to be 3D printed were designed in fusion 360 giving the perfect design parameters so that the parts could be 3D printed efficiently and quickly with minimum supports. The machine has a fixed work bed gantry system where the work bed stays stationary and the X, Y and Z axis moves over this work area. The X and Z axis moves along a single profile, therefore require one carriage each thus one motor required per carriage. The Y axis being the main and the longest axis requires supports from both sides thus requires two carriages with one motor each.

All the electronic control units, power converters, motor drivers, pumps and the micro controllers are placed on the moving vertical beam of the Y axis frame. All the electronic and electrical components are placed here as it reduces the amount of wires and tubes which has to be dragged along the Y axis from the outside. The only connection required from the outside are two wires (+ and -) for power and one tube from the water source. All the wires and tubes runs through drag chains along with supports so that these are protected, organized and tangle free. All the 3D printed parts are printed with 2 mm wall thickness and 70% triangular infill. The material used is PLA. All the 3D printed tools are made up of 2 parts - the tool base which consist of 3 magnets for the attachment to the tool holder and the contact part which comes in contact with the soil. This is done so that in case the tool tip breaks, it can easily be replaced with a new tip instead of replacing the whole tool.

The growing bed is built out of 20mm plywood. It gives a soil depth of 150mm. The soil is prepared by blending red soil with bone powder cow dung and manure. The growing bed has castor wheels attached for easy movement of the setup.

Figure 1. Gantry setup and seed holder with tool bay.
3.2. Network architecture

Object detection is a technology which is related to the image processing and computer vision. It has many applications which include the security and surveillance. Deep neural network - YOLO based object detection algorithm is used for the purpose of detection of plant weed in the image. YOLO can in fact, detect multiple weeds on a single image.

3.2.1. You Only Look Once (YOLO). The YOLO architecture consists of 27 convolution neural network layers, which comprises of 24 convolution layers, 2 Fully Connected layers and a final detection layer.

Figure 2. Position of plant coordinates.

The working area is considered as 840x684 wherein, fixed position coordinates are chosen. It is at these locations, the seed is planted and the plants are grown. Watering is also done at these coordinate locations. For this purpose, it is these coordinate points that are saved in a position file, along with the type of the plant that is planted.

Figure 3. (a) Original YOLO architecture, (b) Image divided into grid, (c) Bounding box predictions and (d) Predictions after non-max suppression.
The input image is divided into grid cells of size $S \times S$. Here, each grid cell predicts $B$ bounding boxes as well as a score for each of these $C$ classes. Each bounding box is said to have 5 predictions. They are $b_x$, by (center x, center y), $w$ (width), $h$ (height) and confidence of the bounding box. For each of grid cell, there will be one set of class scores $C$ for all the bounding boxes. Thus for each image, this tends to result in the output to be a vector of $[S \times S \times (C + 5 \times B)]$ numbers. The features which are extracted from the convolutional layers are used by the fully connected layers. This information is used to predict the object probabilities. The final detection layer is a regression, which maps the final fully connected layer to the bounding box. The original YOLO network trains on PASCAL VOC 2007 and 2012 dataset, which has 20 classes of objects, and a grid size of 7x7.

3.2.2. Dataset. The real images are a major concern as the data on agricultural problems are not easily available [5]. The advancement in the field of farming has not grown as a result of this lack of available data. The main reason being non availability of public plant datasets. Moreover, all the recent achievements are based on the Plant-Village dataset which is available on the Kaggle. Hence, to formulate the dataset as per the requirement, plants of different types were grown and at various stages of the plants, images were captured with different angle, different illumination and different zoom levels. These images were then augmented to provide an increase in the dataset.

3.2.3. Augmentation and annotation. A data augmentation is a way to overcome the shortage of data, which is done by artificially augmenting the data. Larger datasets are required for proper working of Deep learning techniques. But, availability of large dataset is not the case always. Hence, the only way to overcome this shortage of data is by artificially augmenting the dataset. Moreover, issues like over-fitting is better avoided by large dataset. As, the data collection can be a time-consuming process, augmentation technique is used in order to enlarge the datasets. The common method is done by those including pixel wise changes such as adding noise, rotating etc. This adds distortions to the images or in essence to the data. A LabelImg tool is used for the purpose of annotation. Here, each image in the dataset is being annotated and as the result, a YOLO text file is being generated.

3.2.4. Training. Regarding the training, the annotations are reproduced to indicate the bounding box for the ground truth. The bounding box labelling tool is actually used for the purpose of ground truth coordinate creation. The PYTHON GUI is the modified version of labelling software which is created by puzzledqs with the interface. The labelling tool generates four points for the coordinate $(x_1, x_2, y_1, y_2)$ plus the class id. A text file saves the details of the ground truths which are used at the time of training. Nvidia GPU accelerates the training for the network. This is comparatively faster than that by a normal CPU. The weight files are saved after every thousands of iteration. This can be used as a checkpoint in case the training needs to be interrupted and stopped.

3.3. Communication protocol
The communication between the computer and the machine is done by sending string commands from one device to the other. These commands can be to position the machine to a particular coordinate or switch on or off the Vacuum pump and water pump, or pick up or deposit a tool etc. Most of these basic codes are used individually only when the machine is tested during its code and software development and not during daily operations or procedures of the machine.

The whole process of the system will be controlled using Raspberry Pi which provides the actuating signal to the Arduino to do the specified work. The Raspberry Pi sends signal to the Arduino using Serial communication via the USB port. In the program, the required action is entered and a string will be created based on the user entry. The corresponding string is then encoded to bytes and sent via serial port. The data is then read by the Arduino which does the corresponding action. When the work is done it sends a signal back to the raspberry indicating the work is completed.

At the beginning, the user can choose an option among Seeding, Watering, Weed Detection, Showing the Plant bed and Add plant Type. If the user chooses the option Seeding, the program will
ask about the type of plant and the position where the plant to be grown. The program then checks whether the plant is available or not using the plant file (plant.json) and then checks whether the position is available for seeding or if already a plant is grown there with the help of a json file (position.json) which contains the information about the plant types in each position. If a plant is already grown at the point specified it shows an error and asks the user to enter other location. If the specified position is empty, then the program sends the corresponding string to the Arduino.

![Communication protocol](image)

**Figure 4.** Communication protocol.

The next option will be watering the plants. This gives the user two options, water all the plants or water plants in specified position. If the user chooses to water all the plants, then the program checks the position file to find the positions of the plant and the plant grown in those positions. For each plant type, it sends the corresponding signal to the Arduino one after another. If the user chooses to water specified plant, then the positions will be entered by the user and only those locations will be sent to the Arduino. The plant type is obtained from the position file and the corresponding control string is generated.

![Substring command execution](image)

**Figure 5.** Substring command execution.

The next option is to detect the weeds. In order to cover the entire field the camera needs to capture at two images from the field at different positions. The program sends the control signal to the Arduino to move the camera to the specified location for capturing the image. When it reaches the
position, it calls for the program to detect the weeds which returns the position of the weed. These positions are stored in a variable. The program then sends the signal to move the camera to the next position where the image is captured and used to detect the weeds. Once all the weeds are detected, it passes the control signal with the position of the weed.

![Diagram showing the process of weed detection](image)

**Figure 6.** Execution of weed detection.

For the next option to show the plant bed, the program access the position file and check the status of each position whether they are occupied or not. If occupied it shows the annotation for plant type and if not it shows as NA. The last option is to add plant to the plant file. This option is to add new plant details to the plant file which include the plant name along with an annotation and the water required for each plant per day. Once the control string is created, the corresponding string will be encoded to bytes in utf-8 format. These bytes are then passed to Arduino via Serial communication using the USB port. When the actuation is completed, the Arduino sends back a string indicating that the required action has been completed. The program sends the control signal only after it receives the string back from the Arduino.

The details of types of plants that were grown are listed in the table 1. Parameters like time for growth, amount of water required approximately and the interval of watering are considered.

**Table 1.** Details of plants grown.

| Name          | Time for growth (days) | Amount of water required (approx.) (L) | Interval         |
|---------------|------------------------|----------------------------------------|------------------|
| Lady’s finger | 30-50                  | 0.4 – 0.6                              | Daily            |
| Mustard       | 30-40                  | 0.2                                    | Daily            |
| Marigold      | 30-40                  | 0.3                                    | Once in two days |
| Balsam        | 30-40                  | 0.3                                    | Once in two days |
3.4. Results

From the training loss curve in the below figure, it can be seen that the training loss has dropped sharply at the initial stage of training. Also, the loss value converges roughly around 0.18 after the batch iterations of 24000. The performance of the model is evaluated at different batch iterations and the mAP50 is seen to obtain a highest value of 0.83 in 27000 iterations. After this, the mAP starts decreasing as the model tends to overfit.

![Training loss vs. iterations and mAP vs. iterations.](image)

4. Conclusion

The Cartesian coordinate based Farming system is able to plant the seeds, water them and detect the weeds. The user input is sent to the Raspberry Pi using python. The corresponding control signal was created and sent to the Arduino for actuation. The whole unit runs on a 12V power supply. The seeding was done with the help of a vacuum pump and watering was done with the help of 12V water pump. A camera was used to capture image and detect weeds in the plant bed. The detected weeds were removed using a weeder tool that plunges the weed into the soil.

In most small scale home farms, the seeds are sown in a regular pattern manually. The plants are watered on a regular basis manually without considering soil humidity which can sometimes cause water wastage or water shortage. The farmer has to be present physically to water the crops at regular intervals. Most available soil moisture based watering systems works only for a certain spot in which the sensor is placed. Because of this multiple sensors have to be used to cover a larger area. This could be expensive and complicated to maintain. In the busy world of today, people actually do not find time to grow a part of their food even when they have land for it. An automated and simple system like this could encourage people into small scale farming.

This project finds wide applications in backyard, terrace and other small scale farming. It is also helpful when growing different varieties of plants within the same growing area as the machine already knows where it planted each kinds of seed. Therefore the system can care for each kind of plant individually. The main frame of the system is made using standard aluminium profiles, therefore scaling and adapting the system according to the area and shape of land is easier. The weeds that might grow are removed in their infant stages once identified and located by a camera using deep learning.

The proposed system will allow the user to grow crops without the need of much human labor or interference. Only adequate amount of water is provided to the crops therefore it prevents the crops from over-watering and saves water. A system like this encourages people in small scale farming. Uniform sowing of the seeds in the provided area so that the available space is utilized efficiently and conservatively.
5. Future scope

The main future scope of this project is its scalability. With the present setup, the machine can only cover an area of 745 x 990 mm. In this setup, the Y axis can be extended as long as the user wants to cover a longer strip of land just by increasing the length of the Y axis profile, power cables and the inlet water tube. Both the X and Y axis can be extended by using stronger profiles so that a large area can be covered using a single machine. At the same time multiple smaller machines can also be used to cover the same area of land if speed of farming is important. Addition of new tools and procedures for taking care of a variety of difficult to grow plants can be done.

Since the machine works on a voltage range between 12 - 24 volts and the system is idle most of the time other than during the seeding procedures, the system can run on a low energy power source. The whole system can be powered by renewable sources of energy like solar energy. The water source can also be made renewable by rain water harvesting.

The present system knows the exact locations where the crops are planted, therefore all plants detected by the camera outside these locations are considered as the weed and removed. The problem with this method is that weeds growing close to the crops cannot be detected and it is these weeds that affect the crops the most. Therefore in the future models of this farming system, software advancements can be done so that every plant, both weeds and crops, detected by the camera is examined by leaf pattern or shape to determine if it is a crop or weed so that every single weed at every locations can be detected and removed.

References

[1] Cheng B and Matson E T 2015 A feature-based machine learning agent for automatic rice and weed discrimination, Int. Conf. on Artificial Intelligence and Soft Computing, pp. 517–27, Springer, Zakopane, Poland
[2] Elstone L, How K Y, Brodie S, Ghazali M Z, Heath W P and Grieve B 2020 High speed crop and weed identification in lettuce fields for precision weeding Sensors 20 455
[3] Hea Choon Ngo, Ummi Raba’ah Hashim, Yong Wee Sek, Yogan Jaya Kumar and Wan Sing Ke 2019 Weeds detection in agricultural fields using convolutional neural network IJITEE 8
[4] Liakos K G, Busato P, Moshou D, Pearson S and Bochtis D 2018 Machine learning in agriculture: a review. Sensors 18 2674.
[5] Wang A, Zhang W and Wei X 2019 A review on weed detection using ground-based machine vision and image processing techniques Computers and Electronics in Agriculture 158 226–40.
[6] J F Molina, R Gil, C Bojacá, G Díaz and H Franco 2013 Color and size image dataset normalization protocol for natural image classification: A case study in tomato crop pathologies, Symposium of Signals, Images and Artificial Vision
[7] Raja R, Nguyen T T, Slaughter D C and Fennimore S 2020 A real-time weed-crop classification and localisation technique for robotic weed control in lettuce Biosystems Engineering 192 257–74
[8] Rakmi Jauhari, Agung Nugroho Jati and Fairuz Azmi 2017 5th International Conference on Instrumentation, Control, and Automation (ICA) Yogyakarta, Indonesia
[9] Ross Girshik, Jeff Donahue, Trevor Darrell, Jitendra Malik 2016 Region based convolutional networks for accurate object detection and segmentation IEEE
[10] Shaoqing Ren, Kaiming He, Ross Girshik and Jian Sun 2016 Faster RCNN: towards real time object detection with region proposal networks arXiv:1506.01497v3
[11] Joseph Redmon, Santosh Divvala, Ross Girshick and Ali Farhadi 2016 You Only Look Once: unified, real-time object detection arXiv:1506.02640v5