

Article

Low-Cost Robot for Agricultural Image Data Acquisition

Gustavo José Querino Vasconcelos ¹, Gabriel Schubert Ruiz Costa ², Thiago Vallin Spina ¹ and Helio Pedrini ^{1,*} ¹ Institute of Computing, University of Campinas (UNICAMP), Campinas 13083-852, Brazil² The Brazilian Center for Research in Energy and Materials (CNPEM), Campinas 13083-100, Brazil

* Correspondence: helio@ic.unicamp.br

Abstract: More sustainable technologies in agriculture are important not only for increasing crop yields, but also for reducing the use of agrochemicals and improving energy efficiency. Recent advances rely on computer vision systems that differentiate between crops, weeds, and soil. However, manual dataset capture and annotation is labor-intensive, expensive, and time-consuming. Agricultural robots provide many benefits in effectively performing repetitive tasks faster and more accurately than humans, and despite the many advantages of using robots in agriculture, the solutions are still often expensive. In this work, we designed and built a low-cost autonomous robot (DARob) in order to facilitate image acquisition in agricultural fields. The total cost to build the robot was estimated to be around \$850. A low-cost robot to capture datasets in agriculture offers advantages such as affordability, efficiency, accuracy, security, and access to remote areas. Furthermore, we created a new dataset for the segmentation of plants and weeds in bean crops. In total, 228 RGB images with a resolution of 704 × 480 pixels were annotated containing 75.10% soil area, 17.30% crop area and 7.58% weed area. The benchmark results were provided by training the dataset using four different deep learning segmentation models.

Keywords: autonomous robots; agriculture; data acquisition; computer vision



Citation: Vasconcelos, G.J.Q.; Costa, G.S.R.; Spina, T.V.; Pedrini, H. Low-Cost Robot for Agricultural Image Data Acquisition. *Agriculture* **2023**, *13*, 413. <https://doi.org/10.3390/agriculture13020413>

Academic Editors: Jin Yuan, Wei Ji and Qingchun Feng

Received: 31 December 2022

Revised: 5 February 2023

Accepted: 7 February 2023

Published: 10 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, agriculture has been increasingly modernized to address existing problems in agricultural fields, where robots are being used to automate repetitive and tedious tasks for humans, in addition to enabling a more effective application of precision agriculture that uses resources much more efficiently. Advanced robotic systems are used from the process of soil preparation, planting, harvesting, and weed control, to the post-processing of the obtained resources. Although agricultural robots are increasingly common on properties around the world, their use is mainly to directly improve crop productivity, while solutions for scientific applications are still scarce and cost prohibitive for the vast majority of research groups, especially in developing countries [1–4].

The complexity inherent in numerous agricultural tasks; for example, harvesting, weed control, and crop quality analysis, among others; is greatly benefited by the advancement of computer vision and artificial intelligence, due to the greater generalization that this type of technology provides. This offers more robust control solutions for different conditions, such as climate, soil, and vegetation [5–13]. The combination of computer vision with machine learning algorithms in the context of precision agriculture can bring numerous benefits, in addition to the drastic reduction of herbicides; for example, the detailed analysis of the phenotype characteristics of plants by quantitative and individual measurements. For this, a spatial analysis of the images captured by the sensors is necessary. Therefore, the use of semantic segmentation algorithms is of paramount importance for detailing the region of interest.

Semantic segmentation is the task of assigning a class label to every pixel in the image. Figure 1 illustrates an example of a semantic segmentation mask in an agricultural image of the Bean dataset, described in detail in Section 5. All soil class pixels in the image are

set to purple, all plant pixels are set to blue, while weeds are set to green. The colors are random and only serve to visually differentiate the classes.

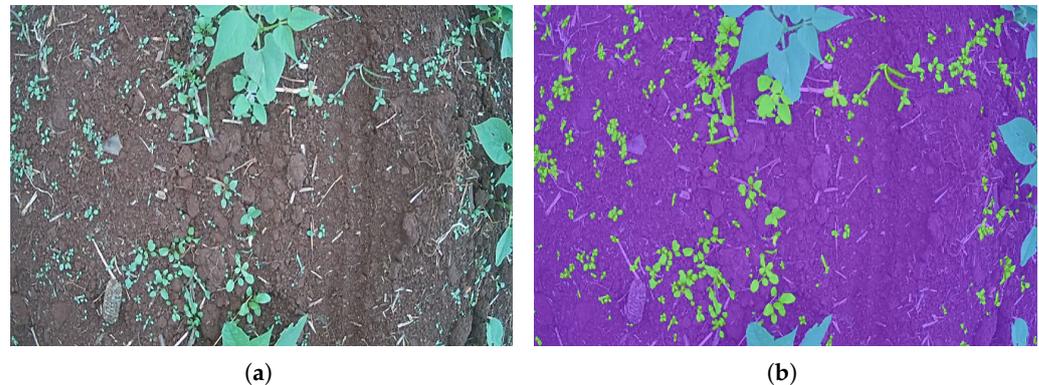


Figure 1. Example of semantic segmentation in an agricultural image. (a) RGB image. (b) Segmentation mask.

Currently, the state of the art in semantic segmentation problems is based on deep neural networks, which require a large amount of annotated data in their training process. Several works present adequate solutions to the weed and crop segmentation problem [5–14]. However, due to the large number of parameters found in this type of model, converging to a robust solution requires the use of a lot of data in its training stage. Therefore, the need for large amounts of annotated images of different plant species and at different growth stages is currently one of the biggest challenges in this area.

However, due to the complex task of image acquisition in the field, such as difficulty in location accuracy, high human effort, difficulty in standardizing capture time, angle, height, and lighting, among others, image datasets for training neural networks are still scarce in the literature, especially for crops of regional interest that are not commonly cultivated on a large scale in Europe and North America. While the popularization of drones is reducing this problem for aerial images, for terrestrial images, which are especially important for active weed control, cost and labor are still major factors in the data scarcity.

Since the vast majority of agricultural robots in the literature are designed to perform more work than just image gathering in the field, when we assess the prices of prototype construction or commercial purchase, we often see prohibitive values, especially in the reality of many universities and research centers in developing nations.

A low-cost robot for capturing datasets in agriculture provides several advantages. First, its low fee makes it more affordable for more farms and research groups to purchase, resulting in broader and more affordable data collection. Furthermore, robots are able to collect data more quickly and accurately than humans, increasing the efficiency of data collection. They also avoid human errors such as fatigue or distraction, ensuring more accurate data collection. In addition, robots can be designed to perform dangerous tasks, reducing the risk to human workers. Finally, robots can be sent to remote areas to collect data where it would be difficult or impossible for humans to reach.

This work aims to present the design, construction and use of a low-cost agricultural robot for image acquisition in agricultural fields. The robot prototype was built with simple, economical, ready-to-use components and was developed for the purpose of acquiring images. In order to provide the community with a new dataset to work with domain adaptation, the designed robot was used to collect a new dataset, called the Bean dataset, which used a crop relevant to Brazilian agriculture; in this case, the bean crop.

2. Terrestrial Agriculture Datasets

This section presents the main datasets used in the area of semantic segmentation of plants and weeds that were acquired terrestrially, that is, close to the ground through robots, or acquired manually. Table 1 illustrates the classes, number of images, resolution

and sensor type of five different datasets present in the literature. It can be observed that only two terrestrial segmentation datasets have more than 60 images, which limits the advancement of deep learning in this field.

Table 1. Comparison of five different terrestrial agriculture segmentation datasets present in the literature.

Name	Classes	# Images	Resolution	Sensor
Sugar Beets [15]	Sugar Beets and 9 weed species	300	1296 × 966	RGB, NIR, Depth and GPS
Sunflower [16]	Sunflower and weeds	500	1296 × 964	RGB
RGBWD [17]	Carrot and weeds	39	3264 × 2248	RGB
CWFID [18]	Carrot and weeds	60	1269 × 966	RGB
CWDD [19]	Carrot, Onion and weeds	40	2419 × 1986	RGB and NIR

The two main terrestrial image datasets available in the literature and used for evaluation of this work are the Sugar Beets and Sunflower datasets. The other datasets contain few images for use in deep learning models.

Sugar Beets is an agricultural dataset composed of 300 RGB images acquired from a terrestrial robot called BoniRob, illustrated in Figure 2, on a sugar beet farm near Bonn, Germany, over a three-month period in the spring of 2016. The images were recorded three times a week, starting with the emergence of plants and stopping when the robot could no longer access the field without damaging the crops. The robot carried a four-channel multispectral camera and a red, green, blue and depth (RGB-D) sensor to capture detailed information about the crop. Additionally, GPS, LIDAR and wheel encoders are available, resulting in around 5TB of data. Only RGB images and their respective semantic segmentation masks from 2–3 weeks of growth stage after emergence were used in this work. An example image found in the Sugar Beets dataset and its respective segmentation mask can be seen in Figure 3.



Figure 2. Acquisition process of the Sugar Beets dataset using the BoniRob terrestrial robot developed by Bosch DeepField Robotics [15]. BoniRob is equipped with four wheels that can be steered independently of each other, allowing flexible movement and navigation on rough terrain.

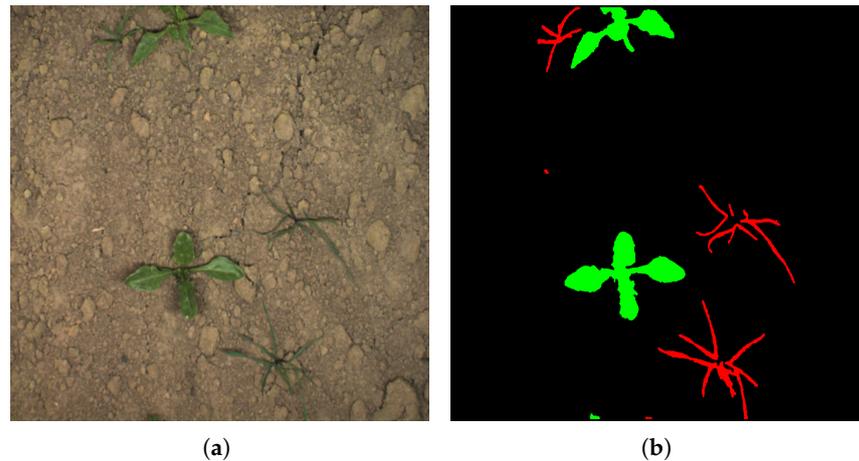


Figure 3. Example of an image found in the Sugar Beets dataset and its respective segmentation mask. (a) Image. (b) Segmentation mask.

Sunflower is a dataset for weed and crop segmentation that was collected using a custom-built agricultural robot, illustrated in Figure 4, on a sunflower farm in Jesi, Italy. The dataset is composed of 500 images and was recorded in the spring season, over a period of one month, starting from the emergence stage of the crop plants and stopping until the end of the useful period. The images were acquired using a four-channel (RGB + NIR) JAI AD-13 camera, mounted on the robot and facing downwards. The dataset provides RGB and NIR images with pixel-wise annotation of three classes: crop, weed and soil. An example image found in the Sunflower dataset and its respective segmentation mask can be seen in Figure 5.



Figure 4. Acquisition process of the Sunflower dataset using a custom-built manual pushcart. Courtesy of Dr. Alberto Pretto—University of Padova.

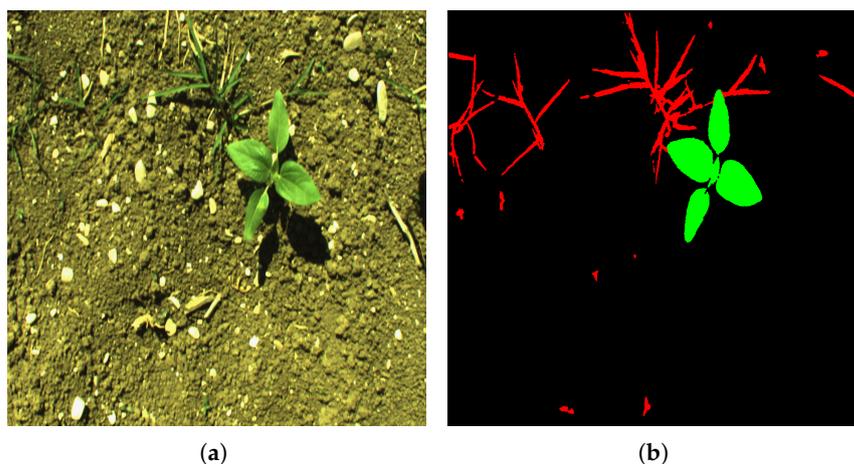


Figure 5. Example of an image found in the Sunflower dataset and its respective segmentation mask. (a) Image. (b) Segmentation mask.

3. Agricultural Robots in the Literature

One of the most important features of agricultural robotics is the replacement of humans with field robots or mechanical systems that can perform dangerous and repetitive tasks more precisely and evenly, at lower cost and with more efficiency. The most common applications for agricultural field robots are weed control and accurate spraying. In this context, spot spraying with robots for weed control application has produced satisfactory results, reducing the use of herbicides to less than 10% of total spraying [20]. Several promising weed robot technologies have been introduced and deployed in recent years as a consequence of multidisciplinary collaboration initiatives involving different academic groups and companies, although they have not yet been fully commercialized.

In this section, we present several previous works that have developed robots to operate in agriculture. However, some of them are not only for image acquisition, but also for chemical or mechanical action for weed manipulation. Despite not being the main focus of the robot presented in this work, it would be possible to modify it to accommodate mechanical actuators. Some of the main agricultural robots in the literature that carry cameras for crop analysis are briefly discussed in this section. Table 2 illustrates the price and country of origin of the robots described in this work. N/A means that no value is available.

Table 2. Prices and places of origin of some common agricultural robots described in the literature.

Robot	Price	Country
BoniRob [15]	€175,000	Germany
Terra-Mepp [21]	\$20,000	EUA
EcoRobotix [22]	\$250,000	Switzerland
Agbot II [23]	€26,000	Australia
Asterix robot [24]	€120,000	Norway
Thorvald II [25]	N/A	Norway
Ladybird [26]	\$1 million	Australia
AgriBOT [27]	N/A	Brazil
DARob (Ours)	\$850	Brazil

BoniRob was developed by students at the University of Osnabrück along with Bosch companies and the German agricultural company Amazone [28]. BoniRob is an agricultural robot that detects weeds using camera technology and image recognition, and then drives a screw into the soil to remove the plant. BoniRob was developed for applications in precision agriculture, that is, for mechanical weed control, selective herbicide spraying, as well as for plant and soil monitoring. It provides assemblies to install different tools for these specific

tasks. BoniRob is equipped with four wheels that can be steered independently of each other, which allows flexible movement and navigation over rough terrain. They provide visual, depth, 3D laser, GPS and odometry data.

The Terra-Mepp robotic platform is designed to semi-autonomously navigate in a single row and view the crop from above, side and below the canopy to provide a comprehensive set of phenotypic data for each plant [21]. An adjustable mast allows flexible placement of sensors to adapt to changes in canopy height throughout the growing season. This system employs an image-based proximal detection approach to measure plant height using stereo cameras and depth sensors. A stereo camera with a 170° wide-angle lens captured the plant height measurements. The camera was mounted vertically at a specific location on the mast, so that the top of the canopy was centered in the camera's field of view. To measure the width of the plant, the low resolution infrared camera was mounted horizontally on the robot's base. Figure 6a illustrates the Terra-Mepp robot.

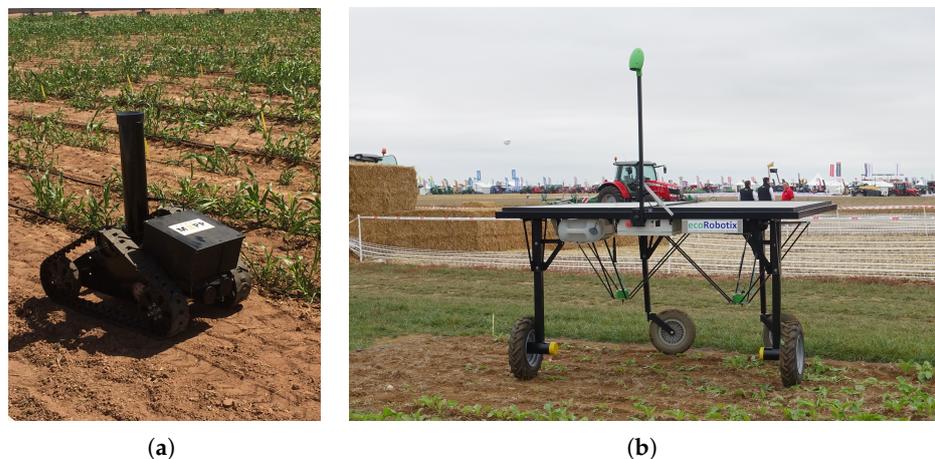


Figure 6. Terra-Mepp (a) and EcoRobotix (b) robots in the field [22,29].

EcoRobotix, illustrated in Figure 6b, is a four-wheeled robot powered by two electric motors, with wheels designed to ride on off-road surfaces, so it could traverse any farmland with relative ease. It has solar panels on top to generate a continuous source of power for the internal battery, allowing it to run as long as there is daylight, thus removing the need to snap and recharge at the end of the day. It weighs approximately 130 kg. An onboard camera, RTK GPS and a series of sensors allow you to identify crops and keep it on a travel course, as well as detect the presence of weeds between crops.

AgBot II is a two-wheel drive (2WD) system agricultural robot developed at Queensland University of Technology for weed and crop management, as well as horticultural applications [23]. The robot is equipped with a downward facing 1.3 MP global shutter camera with an illuminated field of view using a pulsed lighting system synchronized with data capture. The AgBot II can autonomously navigate and traverse a field performing weeding operations, return to the central station to recharge its batteries, and refill its chemical tank before returning to the field to continue operation.

The Asterix robot was designed with the specific task of applying spot spraying herbicides [24]. The robot has a three-wheel design to maintain maneuverability and stability with the benefits of reduced weight, complexity and cost. The vision unit employs a Nvidia Jetson TK1, with a built-in camera unit using Omnivision 4682 4 MP sensor. Raw images are debayered to RGB (Red–Green–Blue) and HSV (Hue–Saturation–Value) color spaces. The forward facing camera and navigation unit allow the tracking of lines across the field. A combination of vision and GPS location detects the end of a row and aids navigation on the headlands.

Thorvald II is a lightweight, battery-powered autonomous agricultural robot developed from the robotics group at the College of Science and Technology, Norwegian University of Life Sciences (NMBU) [25]. The robot is equipped with two drive modules,

two passive wheel modules and no steering modules. As the spacing between the plots is small, approximately 15 cm, the standard wheels of the drive modules are replaced by thinner wheels. The robot is equipped with a RTK-GNSS receiver and an IMU, and navigates through predefined waypoints. It stops and captures images on each plot. The robot is equipped with two pairs of cameras, which consist of an RGB camera and a monochromatic infrared camera. One pair is facing straight down and the other pair is mounted sideways at a 55 degree angle.

The Agricultural Robot (AgriBOT) basically consists of a full-scale four-wheel autonomous robotic tractor equipped with a 4WSD independent steering configuration [27]. The AgriBOT project is a cooperation project between EMBRAPA (Empresa Brasileira de Pesquisa Agropecuária), EESC-USP (Universidade de São Paulo) and Jacto Company. The AgriBOT has two laser sensors attached to its front, pointing downwards at -30° , a model inertial measurement unit (IMU), a global positioning system (GPS), real-time kinematic (RTK) and two red, green, and blue (RGB) cameras with charge-coupled device (CCD) sensors of 1600×1200 pixels of resolution.

Researchers at the Australian Centre for Field Robotics (ACFR) of the University of Sydney have developed the Ladybird robot, which is a lightweight, omnidirectional electric vehicle for advancing agricultural robotics technology. The robot is equipped with a forward and backward facing GPS, Light Detection and Ranging (LIDAR) along with the Point Gray Ladybug 3 spherical camera to capture surroundings data to avoid obstacles and detect crop rows. In addition, the robot has a set of three sensors under the structure. The first is a camera to capture RGB images of the crops, whereas the second is a hyperspectral imaging camera to capture infrared and ultraviolet data. It also has a laser sensor to determine the height of crops above the ground.

Because the vast majority of the aforementioned robots are designed to perform more tasks than just image acquisition in the field, when we analyze the costs of construction of the prototype or for commercial purchase, we observe values that are often prohibitive, especially for the reality of many universities and research centers in underdeveloped countries. As an example, the BoniRob robot used to capture the Bonn dataset costs at around \$250,000.00, while cheaper versions, such as ecoRobotix and AgBot II, cost around \$90,000 and \$26,000.00 at outdated prices. At the other end of the spectrum, robots such as Ladybird have an estimated prototype build value of approximately \$1 million. In addition, we needed an alternative to the pushcart method used in the Sunflower dataset for image capture that did not require as much manual labor. In this context, the contribution of the development and construction of a low-cost terrestrial robot becomes important for the advancement of computer vision science in agriculture.

4. DARob—Data Acquisition Robot

This section presents the design and construction of the DARob robot that aims to capture images in agricultural fields.

The robot was designed to have an unobstructed region at its center, so that there is a gap for plants to pass through as the robot moves along the crop. The design was intended for strip cropping, in which several crops are planted in alternating rows. Typical combinations include corn, sugar beets, soybeans, and grasses such as hay and wheat. This configuration allows the span to be extended as needed for the plantation rows. The distance between the wheels and their height can be modified, thus facilitating the adaptation of the robot to different configurations of the rows, such as paired or twin rows, as well as the adaptation of the height according to the existence of larger plants or in different growth stages.

The robot is 1530 mm long and 635 mm high in its most compact form. The minimum span width and height are 145 mm and 535 mm, respectively, and the maximum span width is 835 mm. The project was also developed so that the robot is fully modular, facilitating its transport and assembly at the site of use. Figure 7 illustrates the DARob. The batteries and

electronic control mechanisms are placed in compartments for easy coupling on the upper part of the robot, which is insulated against dust and moisture.

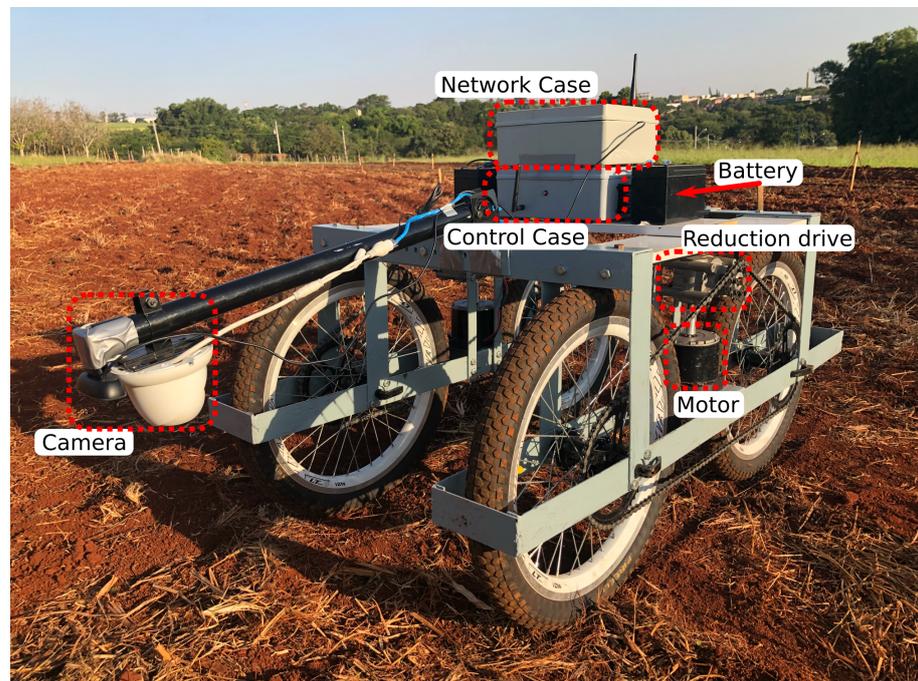


Figure 7. View of the DARob with its main systems.

The robot's maneuvering method is the skid steering type, which is characterized by not having a steering system, but rather changing the direction of the vehicle by controlling the relative speeds of its left and right sides. Although the skid steering mechanism has a simpler construction system, in practice there is a low efficiency for making sharp turns, due to the wheel having to slide sideways to turn. To circumvent this problem, the robot was built to run in either direction (front or back). This means that when the robot reaches the end of the mapped terrain, it can continue driving backwards and does not need to turn 180° to continue, so only minor corrections in the direction are required for it to change planting rows. The robot is controlled by remote control with electric propulsion through the use of two 300 W direct current motors, each controlling the two wheels on its side.

The motors are bolted to the robot's metal frame, and its axes are directly connected with 1:15 transmission reducers, which increases the torque capacity of the motor. A 16-tooth pinion is installed on the output shaft of the gearbox, which is coupled to a chain that connects the 46-tooth chainrings, connected to both wheels (front and rear). This connection also generates a gain in torque of $16/46$, totaling approximately a 43-times gain in the motor's rated torque. The wheels are 52 cm in diameter, and the electric motor has 0.8 Nm of nominal torque and 1.7 Nm in stall.

The mechanical structure was built with the goal of modularity. Therefore, the robot, when disassembled, is separated into four parts of equal size (one for each wheel). However, for transport in smaller places, such as car interiors, simply separating the two sides of the robot is sufficient. All electronic components are also easily removable and separable, as they are placed on a plate above the frame. The entire process of assembling the robot and installing the electronics takes about 10 min and requires only one person.

The camera is placed on a vari-angle bar placed in front of the robot above the space where plants can move. The camera chosen for this project was the Intelbras s4020 IP Camera, a low-cost RGB and infrared camera used mainly for security tasks, which, by construction, is designed to work outdoors in extreme temperatures, weather, and unhealthy environmental conditions.

The camera is designed to handle temperatures ranging from -30 to 50 °C, with a 3.6 mm focal length lens and 1/3" Progressive Scan CMOS sensor. The camera has an IP66 protection rating, making it ideal for working in dusty areas. The transmission of images is carried out via Ethernet cable, its power supply is 12 volts direct current and its consumption is only 4.8 W, with a total weight of 360 g.

The robot's guidance software is controlled by an open hardware platform called Pixhawk, which has magnetometer, accelerometer, barometer, and GPS (Global Positioning System) modules built into the system. These sensors are used as inputs to simultaneous localization and mapping (SLAM), which is calculated from the Ardupilot firmware. Ardupilot is an advanced and reliable open source autopilot software that has been in development since 2010, which from the input of sensors and user commands, controls the current that is delivered to the motors. User control can be performed via remote control or a pre-configured mission via the graphical interface of the QGroundControl ground control station.

QGroundControl is a control station that works with the protocol Micro Air Vehicle Link (MAVLink) compatible with open-source autopilots including the ArduPilot. In this project, QGroundControl was chosen as the ground control station because it provides easy and straightforward use for beginners, as well as offering support to advanced features in mission control and vehicle configurations with autopilot. In addition, QGroundControl is one of the most stable ground control stations, has a simple and efficient interface and is available in different operating systems, such as Windows, Mac OS X, Linux, Android and iOS. Figure 8 illustrates the QGroundControl interface with the waypoints of a planned mission.

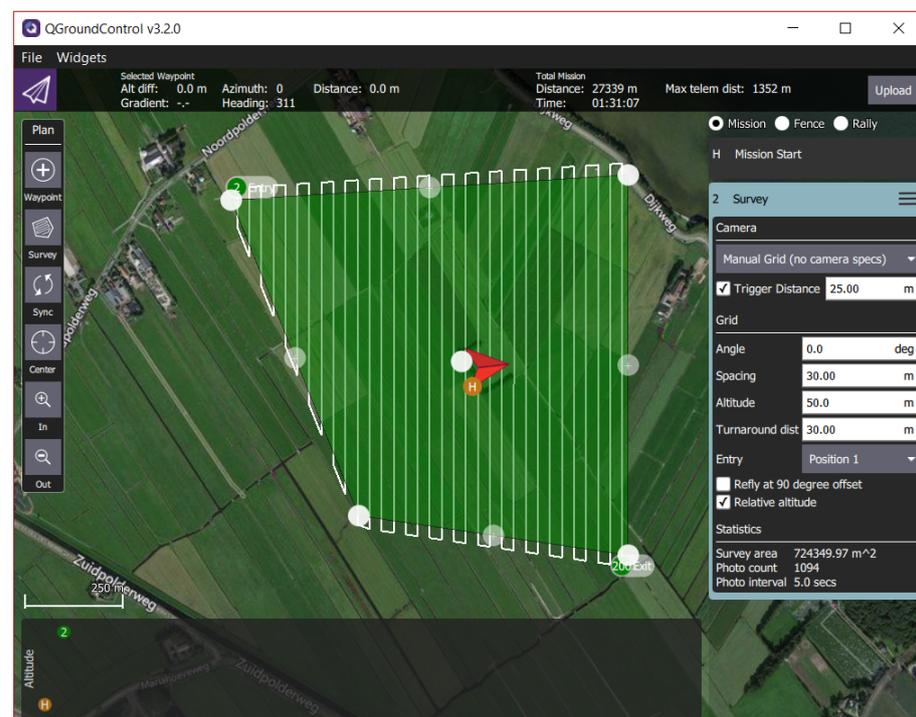


Figure 8. Graphical interface of QGroundControl with an example of the mission's waypoints map. Source: www.qgroundcontrol.com, accessed on 27 March 2022.

In addition to the Pixhawk and the GPS module, the robot's control system also has an H-bridge (model Bts7960) and multiplexers (model CD74HC4067) used to convert the output of the Pixhawk to command the H-bridge. Moreover, the remote control receiver module and the on-off switch button for activating control via Pixhawk are placed in the control system case. Figure 9a illustrates the control system case.

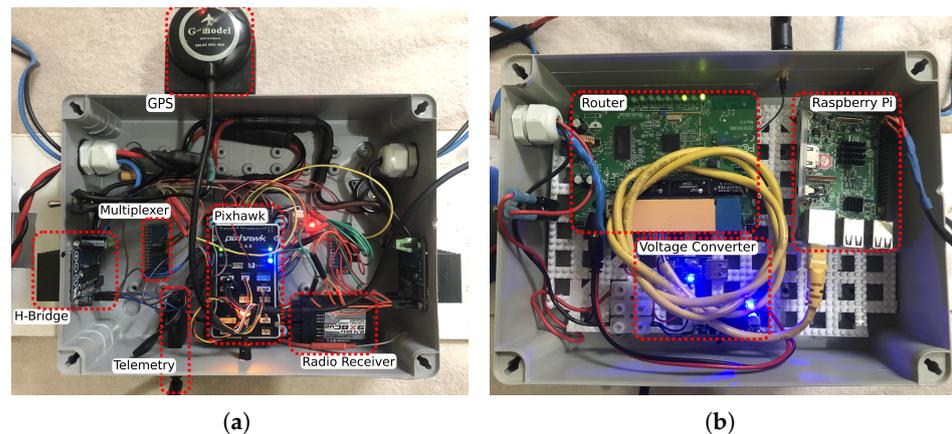


Figure 9. Details of the components present in the control and network systems. (a) Control System. (b) Network System.

In addition to the control system cabinet, the robot has a network system cabinet that includes an embedded microcomputer (Raspberry Pi 3 Model B+) responsible for processing and capturing camera images, and a network router. The images were stored on a 128 GB SD card attached to the Raspberry Pi and, after the missions, the data were transferred to a personal computer. The network router has two different purposes, the first is to allow the raspberry computer to access the IP camera images using a real-time streaming protocol (RTSP), and the second is to allow a ground control computer to access the Raspberry system via Secure Shell (SSH). To carry out this process, the microcomputer runs on Linux (Raspbian OS). The router enables real-time monitoring of collected photos as well as adjusting data collecting settings without having to access or stop the robot. Figure 9b illustrates the network system case. The total cost to build the robot was estimated at around \$850. The main components of the robot and their respective costs are shown in Table 3.

Table 3. Main components of the robot and their respective costs.

Item	Price
Raspbeery Pi 3	\$127
Pixhawk Kit	\$179
Network Router	\$19
2xMotor 300W DC 24V	\$187
2xBts7960 (Driver)	\$13
2xCd74hc4067 (Multiplexer)	\$4
Robot Frame	\$55
Camera Intelbras s4020	\$69
2xBattery 12V	\$39
Reduction Drive	\$99
Extras	\$52
Total	\$843

5. Bean Dataset Acquisition and Annotation Process

From the analysis of Table 1, it can be observed that only two datasets, Sugar Beets and Sunflower, provide a reasonable amount of images for the application of deep neural network models. The amount of data required to train a deep neural network from scratch varies depending on several factors, such as model complexity, data variety, and degree of overfitting. A deep neural network is typically trained with at least hundreds of pictures [30].

In addition, there are few plant variations available in the datasets, and none of them is a plant of great importance for Brazilian agriculture. In this sense, we chose to create a new

dataset to expand the possibilities of evaluating computer vision methods in agriculture and also to create a database more aligned with the interests of Brazilian agriculture, such as plants of the Leguminosae family. Due to the great morphological similarity between the species and the greater availability at the time of planting, we mainly focused on the use of the common bean.

At the end of April 2021, the sowing process of the bean crop began within an experimental area at the School of Agricultural Engineering (FEAGRI) of the University of Campinas (UNICA MP). The area was prepared before sowing by mechanically removing the weeds present in the region. This was the only weed control action in the field, because after the emergence of the plants, no type of control was carried out.

The common bean, the crop used in the proposed dataset, is a perennial herbaceous plant with the scientific name *Phaseolus vulgaris* L. It is a member of the Leguminosae family. It has two different types of leaves: the simple or primary leaves, which are opposite, and the compound leaves, which are composed of three leaflets and arranged alternately (trifoliolate). Beans can be grown for up to four harvests in a year, as they have a short growing cycle (about 90 days). The best sowing time is determined by the common bean's climatic requirements, which are specific to this variety [31].

One difficulty encountered during the capture of the dataset was the modification of the lighting in the field, due to the robot not having protection from direct sunlight. In some cases, images appear with camera shadows, so priority was given to capturing the dataset in the late afternoon when sunlight was not shining directly on the capture area.

We collected data during one month, a specific period during which we covered various plant growth stages. On average, data was acquired once a week, leading to 4 days of capture in total. When recording on a typical day, the robot drove between two rows of cultivation, each measuring approximately 100 m in length. The robot was radio controlled during the data collection process, maintaining an average speed of 10 cm/s and a capture rate of 0.5 frames per second. Figure 10 illustrates DARob in the acquisition process in the FEAGRI field.



Figure 10. DARob in the acquisition process in the FEAGRI field.

In total, more than 5000 images were captured in RGB and IR format. Even though only a subset was annotated with the segmentation maps, all images will be made publicly available. Although the images are captured from emergence to the adult stage of the plant, all of the images annotated in the Bean dataset are of plants in a growth stage no more than

one week apart, 2–3 weeks into the growth stage. Figure 11 illustrates RGB and IR images acquired with DARob. The data collection process covered different growth stages of the bean crop. Our intention was to capture variations over time relevant to weed control. The robot visited various regions of the field in different weather conditions, ranging from sunny to cloudy. Although RGB and IR images were acquired, the applied capture process did not allow a simple capture of the two aligned images.

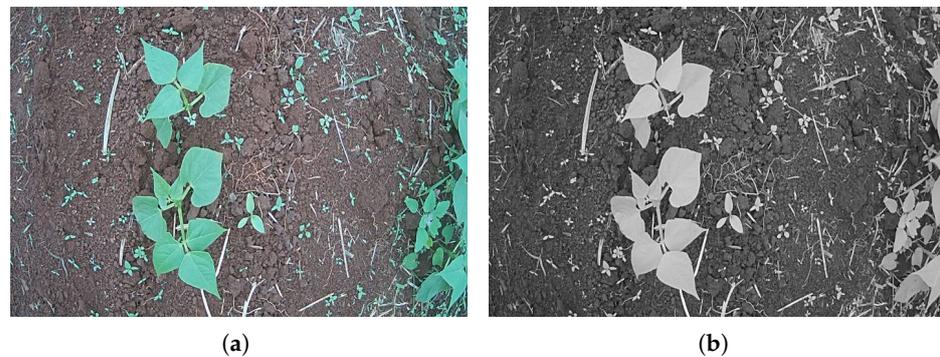


Figure 11. Examples of bean plant images acquired with DARob. (a) RGB image. (b) Infrared image.

Image annotation was performed using the *supervise.ly* platform [32] (version 6.4.31), which is a Web platform for computer vision data annotation. Many annotation tools are available in the *supervise.ly* interface for effective semantic segmentation, including cuboids, polylines, bitmap brushes, bounding boxes, polygons, and keypoints. Figure 12 illustrates the graphical user interface of the platform and the result of a semantic segmentation case of the annotated images. The process was carried out in stages, where the first 10 images were manually annotated, and later the Deeplab-v3 model [33] with pre-trained ResNet 101 backbone in the COCO dataset [34] was adjusted with the annotated images using 224×224 pixel patches. The inferred images were uploaded back to the platform for correction. After 10 corrected images, the process was repeated to increase the inference quality. After three cycles, no further improvement was noticed, so the rest of the images were manually corrected.

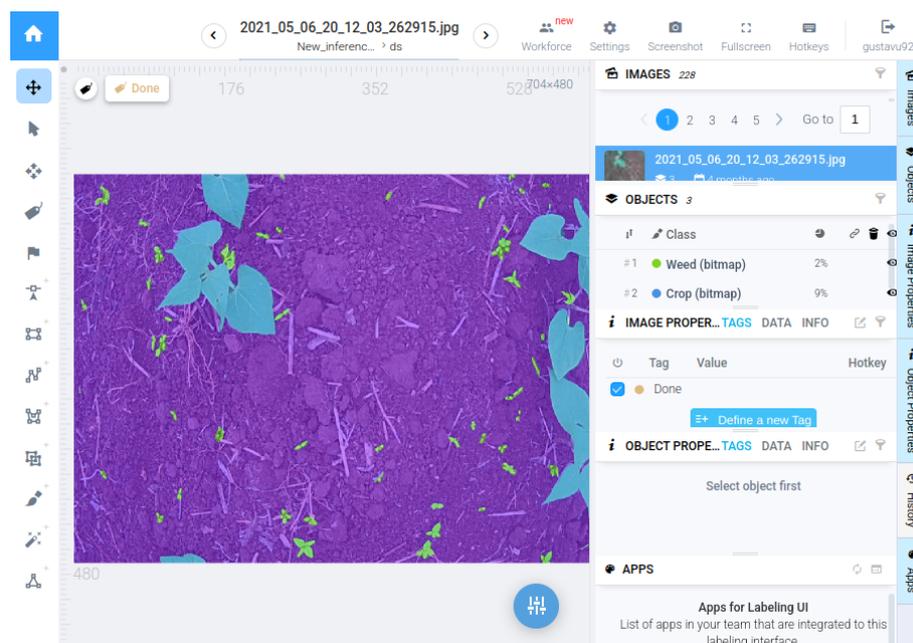


Figure 12. *Supervise.ly* graphical user interface in the process of segmenting an image acquired with DARob.

A total of 228 RGB images with a resolution of 704×480 pixels were annotated. The crop growth stage chosen for the annotation was similar to the steps present in the Sugar Beets and Sunflower datasets to enable validation of the domain transfer models. The available dataset was separated into five k -fold sets, each one with 183 training images and 45 testing images. The dataset contains 75.10% soil area, 17.30% crop area, and 7.58% weed area. The annotation process lasted approximately 3 months, between November 2021 and January 2022. Each image took approximately 2–3 h to be fully manually annotated. With the iterative process, the average time decreased to 30 min. The dataset is available for download at https://github.com/gustavu92/bean_dataset (accessed on 30 January 2023).

Four different segmentation networks were trained on the dataset to provide benchmark results. For this, we used the BiSeNet [35], DuNet [36], Deeplab-v3 [33], Deeplab-v3+ [37] and PSPNet [38] networks. In all cases, except for BiSeNet, the ResNet50 network was used as the backbone, due to the use of this model in benchmark results in the literature. The BiSeNet network was also evaluated using the ResNet18 backbone. The mIoU results for soil, crop and weed classes are shown in Table 4.

Table 4. Benchmark segmentation results for the acquired Bean dataset.

Model	Bean			
	Soil	Crop	Weed	Mean
BiSeNet (ResNet 18) [35]	0.942	0.920	0.625	0.829
DuNet (ResNet 50) [36]	0.950	0.927	0.662	0.846
Deeplab-v3 (ResNet 50) [33]	0.953	0.935	0.680	0.856
Deeplab-v3+ (ResNet 50) [37]	0.957	0.958	0.682	0.866
PSPNet (ResNet 50) [38]	0.959	0.940	0.708	0.869

The computer used for training is equipped with a GeForce RTX 2070 with 8GB of memory, a tenth-generation Intel i7 CPU and 24GB of RAM. The PyTorch framework was used to implement our methods due to its easy debugging feature and open source license. We also used several libraries for Python programming language, such as NumPy for linear algebra operations, SciPy for scientific computing, OpenCV, PIL and Scikit-Image for computer vision and image processing operations.

Table 4 shows that the tested networks have greater difficulty in segmenting the weed class, which is expected, due to its smaller quantity and smaller size compared to the crop and soil classes. The tests also show good consistency in the results, with little variation due to the complexity of the networks used.

The Sunflower and Sugar Beets datasets were also evaluated with the BiSeNet network, using ResNet18 as the backbone. The results are shown in Table 5. It can be seen that the results for the proposed Bean dataset achieved a higher mIoU in the crop and weed classes. This difference can be explained due to the higher proportion of the weed class in the images. However, it is also observed that the results in the soil class were worse than the results obtained in the other data sets. This can also be explained due to the higher proportion of plants, causing the network to make more errors in the soil class. However, the mIoU in the proposed dataset is slightly higher than the compared datasets.

Table 5. Comparison between the Bean dataset, Sunflower dataset, and Sugar Beet dataset using the BiSeNet network with ResNet 18 as the backbone.

Dataset	Soil	Crop	Weed	Mean
Bean Dataset (Ours)	0.942	0.920	0.625	0.829
Sugar Beets	0.993	0.868	0.430	0.764
Sunflower	0.984	0.840	0.565	0.796

Although Tables 4 and 5 give quantitative values about the results, it is difficult to analyze what would be a good value for the mIoU metric in practice for plant and weed

segmentation. This analysis depends on the application in which the result will be used and the specifics of the dataset used. For example, for applications where the result is used for weed control, an average mIoU value of approximately 80 would already be sufficient to detect the main weeds in the analyzed area, but the higher this value, the smaller the size of the detected and segmented weeds.

6. Conclusions

In this work, we designed and constructed a low-cost autonomous robot (DARob) to facilitate the capture of images in agricultural fields. There are some important features to highlight about DARob:

- Low-cost machine: it employed economical and ready-to-use components, which can facilitate the access of other research groups to this type of data acquisition system, increasing the amount of datasets available;
- Automatic operation: the user can program the robot to execute automatically, following the defined mission, which improves the repeatability of the data generated;
- Remote control: it is possible to follow how the data is being acquired during the robot's movement, allowing the operator to correct the acquisition configuration in real time;
- Portability: the robot was designed to be easy to assemble, transport, and also flexible for different types and sizes of crop.

During the operation of the robot, some limiting points were observed:

- Autonomy: the batteries have limited autonomy, reducing the robot's efficiency;
- Bicycle wheels: by using these, the robot has difficulty moving over mud, which reduces the autonomy and disturbs navigability during automatic operation;
- Shadow on images: since it does not have any type of enclosure for the camera, light changing and shadow cause disturbances on the acquired images.

Furthermore, we created a new dataset for segmentation of plants and weeds in bean crops. In total, 228 RGB images with a resolution of 704×480 pixels were annotated containing 75.10% soil area, 17.30% crop area and 7.58% weed area. The benchmark results were provided by training the dataset using four different deep learning segmentation models.

Author Contributions: Data curation, G.J.Q.V. and G.S.R.C.; formal analysis, T.V.S. and H.P.; funding acquisition, H.P.; methodology, G.J.Q.V., G.S.R.C., T.V.S. and H.P.; writing—original draft, G.J.Q.V.; writing—review and editing, T.V.S. and H.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The data presented in this study are available in the main text.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Shah, T.M.; Nasika, D.P.B.; Otterpohl, R. Plant and Weed Identifier Robot as an Agroecological Tool Using Artificial Neural Networks for Image Identification. *Agriculture* **2021**, *11*, 222. [[CrossRef](#)]
2. Ganguly, R.K.; Mukherjee, A.; Chakraborty, S.K.; Verma, J.P. Impact of agrochemical application in sustainable agriculture. In *New and Future Developments in Microbial Biotechnology and Bioengineering*; Elsevier: Amsterdam, The Netherlands, 2021; pp. 15–24.
3. Esposito, M.; Crimaldi, M.; Cirillo, V.; Sarghini, F.; Maggio, A. Drone and Sensor Technology for Sustainable Weed Management: A Review. *Chem. Biol. Technol. Agric.* **2021**, *8*, 1–11. [[CrossRef](#)]
4. Groeneveld, D.; Tekinerdogan, B.; Garousi, V.; Catal, C. A Domain-specific Language Framework for Farm Management Information Systems in Precision Agriculture. *Precis. Agric.* **2021**, *22*, 1067–1106. [[CrossRef](#)]
5. Picon, A.; San-Emeterio, M.G.; Bereciartua-Perez, A.; Klukas, C.; Eggers, T.; Navarra-Mestre, R. Deep Learning-based Segmentation of Multiple Species of Weeds and Corn Crop using Synthetic and Real Image Datasets. *Comput. Electron. Agric.* **2022**, *194*, 106719. [[CrossRef](#)]

6. Mishra, A.M.; Harnal, S.; Gautam, V.; Tiwari, R.; Upadhyay, S. Weed Density Estimation in Soya Bean Crop using Deep Convolutional Neural Networks in Smart Agriculture. *J. Plant Dis. Prot.* **2022**, *129*, 593–604. [CrossRef]
7. Vayssade, J.A.; Jones, G.; Gée, C.; Paoli, J.N. Pixelwise Instance Segmentation of Leaves in Dense Foliage. *Comput. Electron. Agric.* **2022**, *195*, 106797. [CrossRef]
8. Champ, J.; Mora-Fallas, A.; Goëau, H.; Mata-Montero, E.; Bonnet, P.; Joly, A. Instance Segmentation for the Fine Detection of Crop and Weed Plants by Precision Agricultural Robots. *Appl. Plant Sci.* **2020**, *8*, e11373. [CrossRef]
9. Khan, A.; Ilyas, T.; Umraiz, M.; Mannan, Z.I.; Kim, H. CED-Net: Crops and Weeds Segmentation for Smart Farming using a Small Cascaded Encoder-Decoder Architecture. *Electronics* **2020**, *9*, 1602. [CrossRef]
10. Sabzi, S.; Abbaspour-Gilandeh, Y.; Arribas, J.I. An Automatic Visible-Range Video Weed Detection, Segmentation and Classification Prototype in Potato Field. *Heliyon* **2020**, *6*, e03685. [CrossRef]
11. Sheikh, R.; Milioto, A.; Lottes, P.; Stachniss, C.; Bennewitz, M.; Schultz, T. Gradient and Log-based Active Learning for Semantic Segmentation of Crop and Weed for Agricultural Robots. In Proceedings of the IEEE International Conference on Robotics and Automation, Paris, France, 31 May–31 August 2020; pp. 1350–1356.
12. Sodjinou, S.G.; Mohammadi, V.; Mahama, A.T.S.; Gouton, P. A Deep Semantic Segmentation-based Algorithm to Segment Crops and Weeds in Agronomic Color Images. *Inf. Process. Agric.* **2021**, *9*, 355–364. [CrossRef]
13. Zou, K.; Chen, X.; Zhang, F.; Zhou, H.; Zhang, C. A Field Weed Density Evaluation Method Based on UAV Imaging and Modified U-Net. *Remote Sens.* **2021**, *13*, 310. [CrossRef]
14. Vasconcelos, G.; Spina, T.; Pedrini, H. Low-Cost Domain Adaptation for Crop and Weed Segmentation. In Proceedings of the 25th Iberoamerican Congress on Pattern Recognition, Porto, Portugal, 10–13 May 2021; pp. 141–150.
15. Chebrolu, N.; Lottes, P.; Schaefer, A.; Winterhalter, W.; Burgard, W.; Stachniss, C. Agricultural Robot Dataset For Plant Classification, Localization And Mapping On Sugar Beet Fields. *Int. J. Robot. Res.* **2017**, *36*, 1045–1052. [CrossRef]
16. Fawakherji, M.; Potena, C.; Pretto, A.; Bloisi, D.D.; Nardi, D. Multi-Spectral Image Synthesis for Crop/Weed Segmentation in Precision Farming. *Robot. Auton. Syst.* **2021**, *146*, 103861. [CrossRef]
17. Lameski, P.; Zdravevski, E.; Trajkovik, V.; Kulakov, A. Weed Detection Dataset with RGB Images Taken Under Variable Light Conditions. In Proceedings of the ICT Innovations 2017, Skopje, North Macedonia, 18–23 September 2017; Trajanov, D., Bakeva, V., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 112–119.
18. Haug, S.; Ostermann, J. A Crop/Weed Field Image Dataset for the Evaluation of Computer Vision Based Precision Agriculture Tasks. In Proceedings of the European Conference on Computer Vision, Boston, MA, USA, 7–12 June 2015; pp. 105–116.
19. Bosilj, P.; Aptoula, E.; Duckett, T.; Cielniak, G. Transfer Learning between Crop Types for Semantic Segmentation of Crops versus Weeds in Precision Agriculture. *J. Field Robot.* **2019**, *37*, 7–19. [CrossRef]
20. Yeshe, A.; Gourkhede, P.; Vaidya, P. Blue River Technology: Futuristic Approach of Precision Farming. *Just Agric.* **2022**, *2*, 1–14.
21. Kayacan, E.; Kayacan, E.; Ramon, H.; Saeys, W. Towards Agrobots: Identification of the Yaw Dynamics and Trajectory Tracking of an Autonomous Tractor. *Comput. Electron. Agric.* **2015**, *115*, 78–87. [CrossRef]
22. Ecorobotix. The Autonomous Robot Weeder from Ecorobotix. 2017. Available online: <https://www.agriculture-xprt.com/products/autonomous-robot-weeder-579629> (accessed on 10 August 2021).
23. Bawden, O.; Ball, D.; Kulk, J.; Perez, T.; Russell, R. A Lightweight, Modular Robotic Vehicle for the Sustainable Intensification of Agriculture. In Proceedings of the 16th Australasian Conference on Robotics and Automation. Australian Robotics and Automation Association (ARAA), Melbourne, VIC, Australia, 2–4 December 2014; pp. 1–9.
24. Utstumo, T.; Urdal, F.; Brevik, A.; Dørum, J.; Netland, J.; Overskeid, O.; Berge, T.; Gravidahl, J. Robotic In-Row Weed Control in Vegetables. *Comput. Electron. Agric.* **2018**, *154*, 36–45. [CrossRef]
25. Grimstad, L.; From, P.J. Thorvald II: A Modular and Re-configurable Agricultural Robot. *IFAC PapersOnLine* **2017**, *50*, 4588–4593. [CrossRef]
26. Underwood, J.; Wendel, A.; Schofield, B.; McMurray, L.; Kimber, R. Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *J. Field Robot.* **2017**, *34*, 1061–1083. [CrossRef]
27. Diaz, J.F.A.; Guerrero, H.B.; Tiberti, A.J.; Tabile, R.A.; Tangerino, G.T.; Torres, C.J.; Souza, R. AgriBOT-Mobile Robot to Support of Agricultural Precision Activities. In Proceedings of the 22nd International Congress of Mechanical Engineering, Ribeirao Preto, Brasil, 3–7 November 2013; pp. 6243–6254.
28. Biber, P.; Weiss, U.; Dorna, M.; Albert, A. Navigation System of the Autonomous Agricultural Robot BoniRob. In Proceedings of the Workshop on Agricultural Robotics: Enabling Safe, Efficient, and Affordable Robots for Food Production, Vilamoura, Portugal, 11–12 October 2012.
29. Young, S.N.; Kayacan, E.; Peschel, J.M. Design and Field Evaluation of a Ground Robot for High-Throughput Phenotyping of Energy Sorghum. *Precis. Agric.* **2019**, *20*, 697–722. [CrossRef]
30. Cho, J.; Lee, K.; Shin, E.; Choy, G.; Do, S. How Much Data is Needed to Train a Medical Image Deep Learning System to Achieve Necessary High Accuracy? *arXiv* **2015**, arXiv:1511.06348.
31. Karavidas, I.; Ntatsi, G.; Vougeleka, V.; Karkanis, A.; Ntanasi, T.; Saitanis, C.; Agathokleous, E.; Ropokis, A.; Sabatino, L.; Tran, F. Agronomic Practices to Increase the Yield and Quality of Common Bean (*Phaseolus vulgaris* L.): A Systematic Review. *Agronomy* **2022**, *12*, 271. [CrossRef]
32. Supervisely—Web Platform for Computer Vision. Annotation, Training and Deploy. 2022. Available online: <https://supervisely/> (accessed on 28 March 2022).

33. Chen, L.; Papandreou, G.; Schroff, F.; Adam, H. Rethinking Atrous Convolution for Semantic Image Segmentation. *arXiv* **2017**, arXiv:1706.05587.
34. Lin, T.Y.; Maire, M.; Belongie, S.; Hays, J.; Perona, P.; Ramanan, D.; Dollár, P.; Zitnick, C.L. Microsoft COCO: Common Objects in Context. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014; pp. 740–755.
35. Yu, C.; Wang, J.; Peng, C.; Gao, C.; Yu, G.; Sang, N. BiSeNet: Bilateral Segmentation Network for Real-time Semantic Segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
36. Jin, Q.; Meng, Z.; Pham, T.D.; Chen, Q.; Wei, L.; Su, R. DUNet: A Deformable Network for Retinal Vessel Segmentation. *Knowl.-Based Syst.* **2019**, *178*, 149–162. [[CrossRef](#)]
37. Chen, L.; Zhu, Y.; Papandreou, G.; Schroff, F.; Adam, H. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018.
38. Zhao, H.; Shi, J.; Qi, X.; Wang, X.; Jia, J. Pyramid Scene Parsing Network. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.