Multi Species Weed Detection with RetinaNet One-Step Network in Maize Field

J.M. López Correa[1,2], M. Todeschini[1], D.S. Pérez[2], J. Karouta[1], F. Bromberg[1], Ribeiro A, and D. Andújar[1]

[1]*CSIC, Centre for Automation and Robotics,Spanish National Research Council, CSIC-UPM, Ctra. M300 Campo Real, Km 0,200, Arganda del Rey - 28500 Madrid, Spain*
[2]*DHARMa, Artificial Intelligence Laboratory, UTN, Coronel Rodríguez 273, Ciudad - Capital, M5500 Mendoza, Argentina.*
jlopezcorrea@conicet.gov.ar

Abstract

Weed density and composition are not uniform throughout the field, nevertheless, the conventional approach is to carry out a uniform application. *Object Detection Networks* have already arrived in agricultural applications that can be used for weed management. The current study developed a detection and classification weeds system in a one-step procedure using *RetinaNet Object Detection Network*. The procedure was based on identifying *Solanum nigrum L., Cyperus rotundus L.* and *Echinochloa crus-galli L.* and two growth stages both for a broadleaf species (*Solanum nigrum L.*) as well as narrow-leaved species (*Cyperus rotundus L.*) in the maize field. The predictions were evaluated by mAP metric. The result obtained was 0,88 with values between 0,98 and 0,75 depending on the class.

Keywords: Deep Learning, Object Detection Networks, RetinaNet, SSWM, DACWEED.

Introduction

Weed management plays one of the most important roles in agriculture. The weed density and composition are, in practice, not uniform throughout the field, with spatial and temporal variation (Fernández-Quintanilla et al 2018). Until recently the conventional approach was to create a uniform application (Pérez-Ortiz et al., 2016). With *Site-Specific Weed Management* (SSWM), a more precise weed treatment can be achieved with both economic, environmental and food quality benefits (Tang, et al. 2017). SSWM is achieved by applying a treatment only on the weed patches (eg. nozzles on/off), potentially complemented with selective application. An application system consists of a weed mapping component that determines the presence or absence of weeds/crop in the different locations of the field. Based on that information, the implement should adjust its treatment accordingly. Integrating a multi-class weed detection system has the potential of further benefits, with reduced herbicide usage, safer agricultural products for the consumer and the environment, and a more sustainable agriculture. The main expected outcome and impact of this activity is to reduce the herbicide input, herbicide residues in the food chain and costs of weed control.

The use of deep learning has swiftly emerged as a promising method in weed and plant classification and detection, with promising potential for multi-class weed detection problems. Dyrmann et al. (2016) recompiled six datasets from different research works and trained a Convolutional Neural Network (CNN) to classify between twenty two weed species. In this work, a classification accuracy of between 82.4% and 88.2% was achieved showing potential for CNN weed classification leaving aside the problem of the weed detection. Then, Dyrmann et al. (2017) worked on the detection and classification of seventeen weed species in the maize field. Segmentation techniques were used for detección the species plants with the intersection-over-union values of between 0.69 and 0.93 for weeds and maize plants. Following the detection of the weeds, a convolutional neural network was used, which classifies the weeds with an overall accuracy of 87%. However two different processes were needed to identify the weed species with classification accuracy similar to those reported in this paper for detection and classification in one-step.

Lin et al. (2017) developed a new *Object detection Network* called RetinaNet. This network is a one-stage object detector that can handle object location and classification at once, reducing the prediction time and increasing the detection accuracy. Agricultural applications of object detection are becoming more common. A real time vegetable detection system was developed using deep learning networks (Zheng et al., 2018). Multiple object detectors were selected in order to recognize tomato and cucumber at different stages. Among all advanced detectors selected, YOLOv2 (Redmon & Farhadi, 2017) had the highest performance to find all relevant objects. By optimizing YOLOv2, Koirala et al. (2019) developed a new algorithm which exhibited improved detection performance to detect mangos in images. Quan et al., (2019) trained a faster R-CNN network to detect maize seedling under different growth stages and complex field environments. Nevertheless, these methods have not been extensively tested in weed detection. With regards to RetinaNet there are only a few studies. Rançon et al., (2019) tested it in vineyards to detect Esca disease (*Phaeoacremonium aleophilum*) and obtained AP of 0,7. Wei et al., (2020) created a robot "ROS-Based Rapid Identification Robot" to detect weeds in the maize crop with 0,94 AP in weeds and 0,93 maize class. In all cases, the authors have demonstrated the potential use of object detection architectures in the agricultural discipline. Nevertheless, they could not detect and discriminate between several weed species in one-step, like as RetinaNet does. The use of these neural networks algorithms is very useful in today's agricultural conditions. The *Object Detection Networks* can be useful through apps for decision making. On the other hand, these networks can be integrated into real-time treatment systems capable of reducing the aggressiveness of treatments. Some commercial prototypes, such as the Blueriver case, combine real-time sensing systems with a commercial plant-by-plant applicator. All these systems are evolving rapidly and can be a response to the demand for reduced pesticide use and increased yields. The current study proposes a detection system of the weed species *Cyperus rotundus L., Echinochloa crus-galli L., Solanum nigrum L.* and two growth stages both for a broadleaf species (*Solanum nigrum L.)* as well as narrow-leaved species (*Cyperus rotundus L.*); combined with crop plants (*Zea Mays L.*) under real and commercial production fields. The detection and classification of weed species and growth stages were done through a one-step procedure using RetinaNet Object Detection Network.

Materials and methods

Image acquisition

The selected fields for the experiments were located in the province of Badajoz, Spain (39° 1' 14,42", -6° 3' 40,69"). All the images were collected on commercial fields under real uncontrolled illumination conditions. The images were captured in April of 2020 in maize crops (*Zea mays L.*). Two series with different sowing dates were selected such that the crops corresponded with V1 and V3 growth stages (Ritchie et al., 1993). No weed treatment was carried out during the acquisition phase or 10 days before. Each lot was processed following an "M" trajectory, and every 2 m a zenithally image was taken from a height of 1.3 m. Three weed species (*Cyperus rotundus L., Echinochloa crus-galli L., and Solanum nigrum L.*) were found in low growth stages (Vc and or greater or equal to V1), combined with crop plants (*Zea Mays L.*). The images were captured using a Canon PowerShot SX540 HS camera using a resolution of 5184 px X 3886 px. A shutter speed of 1/1000 was used, while the ISO calibrated automatically to achieve a good image quality under the changing lighting conditions during the measurements. Crop and weed plants were in the scene of the image, as well also weeds were in the line and interline crops. Large proportion of weeds in the crop line overlapped with maize plants, especially maize plants in V3 growth stages. Some examples of the dataset are shown in Figure 1.

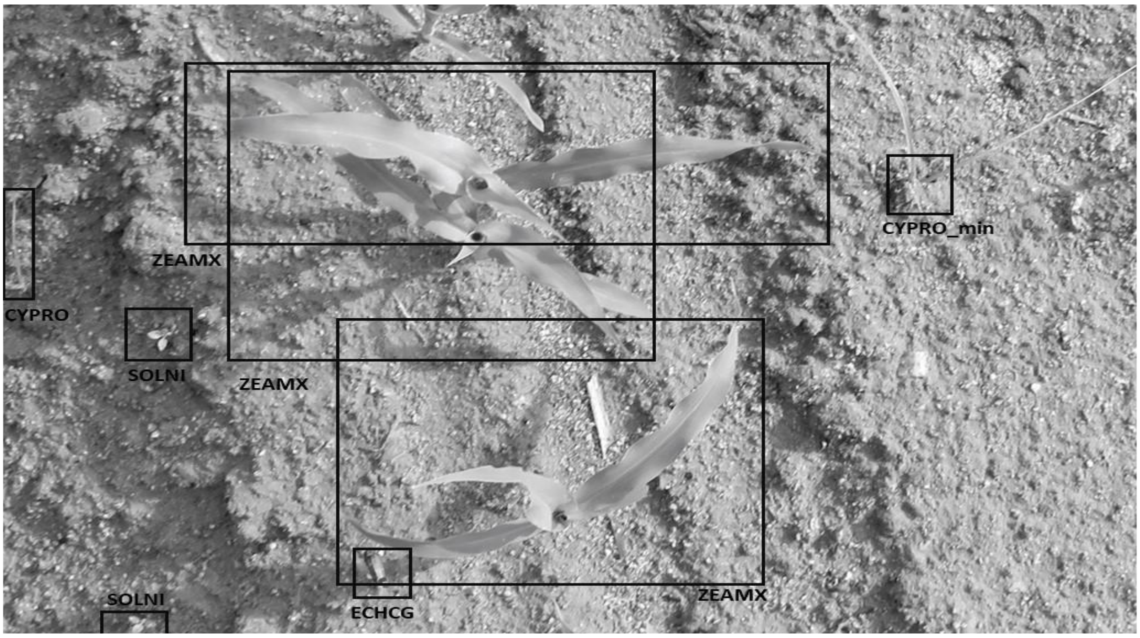

Figure 1: Part of a training image showing the plants classes throughout the bounding boxes and EPPO code.

Image pre-processing

A total of 312 images were taken, in which the present plant species were identified and manually labelled. Besides the previously defined classes, an extra "Ve" class was introduced, containing all emerging plants that could not be identified due to their small size. Labeling the images was done using dedicated software called LabelImg (Tzutalin, 2015), which saves the labels using the PascalVOC format (Everingham et al., 2010). This format saves the co-ordinates of a bounding box surrounding the object that it represents. All bounding boxes defined for one image were saved in a single XML file, and for each object saved the co-ordinates of the upper-left and lower-right corner (xmin, ymin, xmax, ymax), as well as the class name of the labelled object. Four plant species were categorized with the help of experts and labelled using the EPPO code (EPPO code system,), extended using a suffix representing either the growth stage or another identifier as explained below: SOLNI_Vc for plants identified as the species *Solanum nigrum L*, where suffix "Vc" stands for the embryonic growth stage (cotyledon); SOLNI for plants identified as the species *Solanum nigrum L.* showing more leaves than only the embryonic ones; CYPRO_min for plants identified as the

species *Cyperus rotundus L.* showing more than one leaf, where suffix "min" means that for this class the bounding boxes were drawn encompassing only of the centre of the plant (see Figure 1); CYPRO for plants identified as the species *Cyperus rotundus L.* showing only one leaf (>V1); ECHCG for plants identified as the species *Echinochloa crus galli L*; and ZEAMX for plants identified as the species *Zea mays L*.
In order to evaluate the predictive potential of a model, it is necessary to perform these evaluations with data that have not been used for training, i.e., 'new data for the model'. The dataset was divided as follows: 70% of the images were randomly taken to form the *Training Set*, with the remaining 30% being the *Validation Set*. In addition, a random sample of 30% of the images was taken from the training set, conforming the *Test Set*. The last set was taken to evaluate the training progress. Both training and validation (including testing) was conducted using a GeForce GTX 1080 GPU. Captured image sizes could not be processed with the processing power of the GPU, requiring a reduction of their size. For that, a scan of the images was performed generating 74 smaller sub-images for each full-image. Three parameters had to be selected: image height, image width and overlap between one sub-mage and the next; choosing 3886 px for the width, 1926 px for the height, and an overlap of 1900 px. Each XML that defined the plant labels in the original images were corrected accordingly, resulting in a new XML file for each sub image. As a result of the scanning process, some bounding boxes that are in contact with edges of the sub-image might be cut, resulting in incomplete data. To avoid the use of this data for training, those boxes were eliminated. The large overlap chosen for scanning allows that if a plant is removed because it touches the edge it will appear in the next sub-image, thereby ensuring that all the plants in the original images have a label on the training set. After the processing of scanning each of the three image groups (training set, test set and validation set), 4368 training sub-images, 2181 validation sub-images and 1310 test sub-images were obtained. From these, any sub-images without bounding boxes were deleted. The labels were visually examined by experts who corrected labelling errors. The final data set consists of 7859 sub-images with 60436 bounding boxes.

Table 1: Identified weed species in the study labelled with the EPPO code. Number of labels boxes in total for the *Training set, Test set* and *Validation set*.

Species	Label	Training set	Test set	Validation set t
Solanum nigrum Vc	SOLNI_Vc	6836	2025	3469
Solanum nigrum >Vc	SOLNI	2714	907	1275
Echinochloa crus galli	ECHCG	2463	851	1187
Cyperus rotundus	CYPRO	6292	1942	3304
Cyperus rotundus >V1	CYPRO_min	5000	1445	2501
Zea mays	ZEAMX	10193	3004	5028

RetinaNet object detection

Object detection is a powerful technique that can achieve successful plant and weed identification. RetinaNet is a single, unified network composed of a backbone network and two task-specific subnetworks. The backbone is responsible for computing a convolutional feature map over an entire input image and is an off-the-self convolutional network. The first subnet performs convolutional object classification on the backbone's output; the second subnet performs convolutional bounding box regression (Lin et al., 2017). The two subnetworks feature a simple design for one-stage, dense detection. RetinaNet was selected among the object detection networks available in the literature as it has shown a good performance achieving accuracy similar to two-step networks but with the speed of single-step networks.

<u>Training</u>

Wee and crop plants were classified in 6 different classes (SOLNI_Vc, SOLNI, CYPRO, CYPRO_min, ECHCG, ZEAMX). Training was made using the implementation proposed by Gaiser et. al., (2019). The deep learning model used in this study is implemented using Keras 2.4.3 in python 3.6.8 with the TensorFlow (2.3.0) backend.

In most deep learning applications, it is common to utilize a pre-developed computer vision model trained on a relevant dataset (so-called transfer learning). Collecting a large enough dataset for developing a custom deep learning method would be difficult, time-consuming, and nearly impossible for most users focused on application. Using transfer learning (Pan & Yang., 2010), existing feature extraction methods, such as those mentioned previously, can be leveraged from models trained on standard datasets and object detection is fine-tuned to the desired target (Kamilaris & Prenafeta-Boldú, 2018). The Resnet50 (He et al., 2016) model was used as the back-bone network, pre-trained using the COCO dataset (Lin et al., 2014), and an initial learning rate of 1e-5 was used in training. For the training and validation subsets, data augmentation (Shorten et al., 2019) was also undertaken to avoid overfitting and overcome the highly variable nature of the target classification. The Keras library (Keras, 2015) was used to perfom the data augmentation parameter like rotation, scale, illumination, perspective and colour. Specifically, rotation of up to 10º, a brightness shift of ±20 %, a channel shift of ±30 %, a zoom of ±20 % were randomly undertaken, along with possible horizontal and vertical flips. RetinaNet was trained until the mAP of the training set of every class did not improve for 16 consecutive epochs. The parameters used for the implementation of RetinaNet (Gaiser et al., 2019) were the following: *1e-5* for the learning rate; the COCO dataset as pre-trained model, the backbone model used was ResNet-50 (He et al., 2016)), the number of epochs to train was 100 and the number of steps per epoch was 600, the backbone layers were frozen during training, the images were not resized, option image augmentation carried out as stated above, and the batch size was 8. The rest of parameters available in the implementation were left by default.

<u>Fitness Evaluation</u>

A trained model was saved at the end of each epoch. The number of generated models is therefore equal to the number of epochs during the training process in which each trained model may produce a different degree of detection accuracy. In this study, 100 models were generated after the training. Each model was analysed to determine which provides the best result. A validation set with a total of 2181 sub-images with 16764 labelled plants (see Table 1) not used in training were used in validation to select the best performing network model developed in this study.

Results

This study considered five different classes to identify: 3 weed species (*Cyperus rotundus L., Echinochloa crus-galli L., Solanum nigrum L.,*) with two classes for two

growth stages of the species *Solanum Nigrum L.*, plus one crop class (*Zea Mays L.*) . The performance of the RetinaNet modeSl was evaluated with the mean average precision (mAP), computed as the average precision (**AP**) over each class (Padilla et al., 2020).
The curves learning rates, and the mAP values for each epoch are shown in Figure 1 for the test set (sign squares) and validation set (sign rhombuses). In epoch 84, the trained model converges and obtains its maximum prediction value (maP: 0,8768) over the *validation set*, see Figure 2. The AP values per class are shown in Table 2, where the lowest value AP occurs for the CYPRO_min class (AP: 0,7492) and the highest for the ZEAMX class (AP: 0,9744). In addition, the prediction time of the 2181 validation images (3886 px width and 1926 height) was 0,2354 second per image.

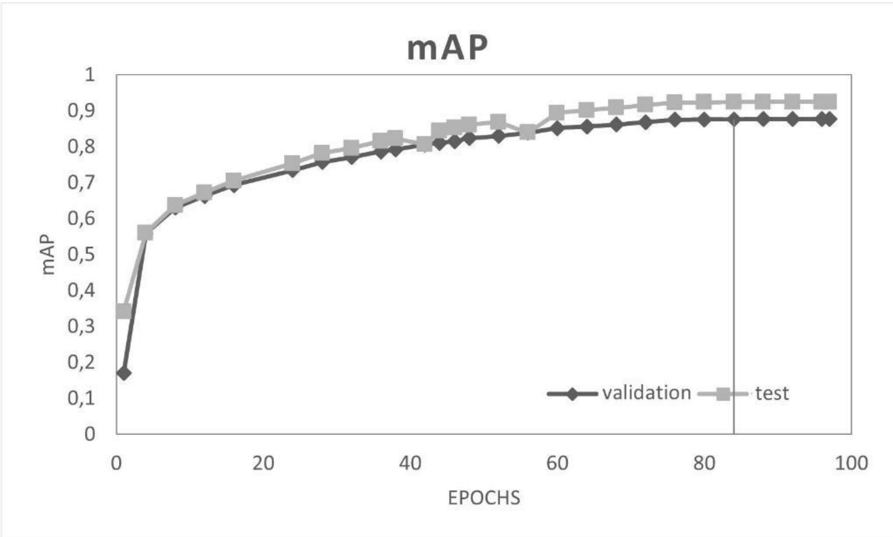

Figure 2: Testing and validation values mAP per epoch and the maximum prediction value for the validation set is shown by the grey line in the 84 epoch.

Table 2: AP per class values and mAP values to validation images with model obtained in 84 epoch.

Class	Label	AP
Solanum nigrum L. Vc	SOLNI_Vc	0,8161
Solanum nigrum L. >Vc	SOLNI	0,9112
Echinochloa crus galli L.	ECHCG	0,9254
Cyperus rotundus L. > 1 leaf	CYPRO	0,8844
Cyperus rotundus L. = 1 leaf	CYPRO_min	0,7492
Zea mays	ZEAMX	0,9744
mAP		**0,8764**

Discussion

The number of species and growth stages of the weeds are low. However the two important groups of weeds were represented, the broadleaf weeds by *Solanum nigrum L.* and narrow-leaved weeds by *Cyperus rotundus L.* and *Echinochloa crus-galli L.* Classification between these groups is important to select the herbicide type because many of these are formulated to control the group of broadleaf weeds, "*broadleaf weeds herbicides*", and others to control the narrow-leaved weeds "*graminicide herbicides*". In addition, the dataset contains an example species for each group (*Solanum nigrum L.* and *Cyperus rotundus L.*) with different growth states. Being able to detect small weeds and low stages of development such as CYPRO_V1 and SOLINI_Vc classes allows early weed control. Therefore, lower dose of herbicide in chemical control or less soil removal in mechanical control. Also, two species (*Echinochloa crus-galli L. Cyperus rotundus L*) of the same botanical family (*Poaceae*) represent morphological similarities for the classification algorithms. Discriminating between these two species is relevant since genetic resistance cases of the chemical family's herbicides have been reported for biotypes of *Echinochloa crus-galli L.* (Gavilan A., 2011). On the other hand, the new self-labeling algorithms are being developed for the research team to increase the number of images. This work includes more weed species and growth stages as well as increasing the amount of data per class. Also, this data repository is being prepared with labeled image and python scripts to leave as open source in the GitHub website. Regarding network configuration, analysis of both performance and inference rate of other Object Detection networks is also planned. Furthermore, cross validation is needed on the parameters of the RetinaNet implementation used to find the best combination.

Conclusions

The current paper demonstrates detecting and classifying in one-step two important groups of weeds, such as broadleaf weeds (*Solanum nigrum L.*) and narrow-leaved weeds (*Cyperus rotundus L., Echinochloa crus-galli L.*) in the maize (*Zea mays L.*) field. In addition, these last two species of the same botanical family (Poaceae) were well classified. Besides that, The *RetinaNet* archives discriminate between two growth stages both for a broadleaf species (*Solanum nigrum L.*) as well as narrow-leaved species (*Cyperus rotundus L.*). This *Object Detection* method demonstrates promising results for porting this knowledge to SSWM not only at the level of weed species but also at their growth state.

Acknowledgements

This research was developed in the "Center for Automation and Robotics of the Universidad Politecnica de Madrid and the Higher Council for Scientific Research" in Spain within the DACWEED project. On the whole with the "Artificial Intelligence Laboratory DHARMa", financed by three CONICET salaries as well as two grants from the Universidad Tecnológica Nacional, and one from the ANPCyT of Argentina.

References

Dyrmann, M., Karstoft, H., & Midtiby, H. S. (2016). Plant species classification using deep convolutional neural networks. Biosystems Engineering, 151, 72-80.

Dyrmann, M. (2017). Automatic detection and classification of weed seedlings under natural light conditions (Doctoral dissertation, University of Southern Denmark).

EPPO code system, http://eppt.eppo.org/, last accessed 1 December 2020; EPPO Plant Protection Thesaurus.

Everingham, M., Van Gool, L., Williams, C. K., Winn, J., & Zisserman, A. 2010. The pascal visual object classes (voc) challenge. International Journal of Computer Vision, 88(2), 303-338.

Fernández-Quintanilla, C.; Peña, J.M.; Andújar, D.; Dorado, J.; Ribeiro, A.; López-Granados, F. 2018. Is the current state of the art of weed monitoring suitable for site-specific weed management in arable crops?, Weed Research, 58, 259–272.

Gaiser, H., de Vries, M., Lacatusu, V., & Williamson, A. 2019. fizyr/keras-retinanet 0.5. 1. Zenodo.

He, K., Zhang, X., Ren, S., & Sun, J. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and pattern recognition (pp. 770-778).

Kamilaris, A.; Prenafeta-Boldú, F.X. 2018. Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70–90.

Keras https://keras.io/. 2015.

Koirala, A., Walsh, K. B., Wang, Z., & McCarthy, C. 2019. Deep learning for real-time fruit detection and orchard fruit load estimation: Benchmarking of 'MangoYOLO'. Precision Agriculture, 20(6), 1107-1135.

Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D. et., 2014. Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.

Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. 2017. Focal loss for dense object detection. In Proceedings of the IEEE international conference on computer vision (pp. 2980-2988).

Padilla, R., Netto, S. L., & da Silva, E. A. 2020. A survey on performance metrics for object-detection algorithms. In 2020 International Conference on Systems, Signals and Image Processing (IWSSIP) (pp. 237-242). IEEE.

Pan, S.J.; Yang, Q. A Survey on Transfer Learning. IEEE Transactions on Knowledge and Data Engineering. 2010. 22, 1345–1359. doi:10.1109/tkde.2009.191.

Pérez-Ortiz, M.; Peña, J.M.; Gutiérrez, P.A.; Torres-Sánchez, J.; Hervás-Martínez, C.; López-Granados, F. 2016. Selecting patterns and features for between- and within-crop-row weed mapping using UAV-imagery. Expert Systems with Applications, 47, 85–94.

Quan, L., Feng, H., Lv, Y., Wang, Q., Zhang, C., Liu, J., & Yuan, Z. 2019. Maize seedling detection under different growth stages and complex field environments based on an improved Faster R–CNN. Biosystems Engineering, 184, 1-23.

Rançon, F., Bombrun, L., Keresztes, B., & Germain, C. 2019. Comparison of sift encoded and deep learning features for the classification and detection of Esca disease in bordeaux vineyards. Remote Sensing, 11(1), 1.

Redmon, J., & Farhadi, A. 2017. YOLO9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7263-7271).

Ritchie, S. W., Hanway, J. J., & Benson, G. O. 1993. How a corn plant develops. Iowa State Univ. Coop. Ext. Serv. Spec. Rep, 48, 21.

Shorten, C., & Khoshgoftaar, T. M. 2019. A survey on image data augmentation for deep learning. Journal of Big Data, 6(1), 60.

Tang, J. L., Chen, X. Q., Miao, R. H., & Wang, D. 2016. Weed detection using image processing under different illumination for site-specific areas spraying. Computers and Electronics in Agriculture, 122, 103-111.

Tzutalin, D. 2015. LabelImg. Git code.
Wei, Q., Peng, J., Zhang, H., Mo, H., & Qin, Y. 2020, October. Design and Implementation of ROS-Based Rapid Identification Robot System. In International Conference on Computer Engineering and Networks (pp. 376-387). Springer, Singapore.
Zheng, Y. Y., Kong, J. L., Jin, X. B., Su, T. L., Nie, M. J., & Bai, Y. T. 2018. Real-Time Vegetables Recognition System based on Deep Learning Network for Agricultural Robots. In 2018 Chinese Automation Congress (CAC) (pp. 2223-2228). IEEE.