A Deep Learning Artichoke Plants Identification Approach for Site-specific UAV Spraying

J. Motta[1], A. Sassu[2], A. Carlevaro[1,3], A. Deidda[2], L. Ghiani[2], F. Gambella[2], G. Garibotto[1]

[1]Aitek SpA, Funded Research Department, Via della Crocetta 15, 16122 Genova, Italy

[2]Department of Agricultural Sciences, University of Sassari, Viale Italia 39 a, 07100 Sassari, Italy

[3]DITEN Department, University of Genoa, Via all'Opera Pia, 11 a, 16145, Genova, Italy

Abstract

The use of Unmanned Air Systems (UASs) in combination with Deep Learning techniques is leading to significant improvements in the management practices of agricultural systems, as evidenced by this study on precise identification and the georeferencing of individual artichoke plants for the development of an on-the-fly UASs aerial spray treatment system.

Data Acquisition

Remote image acquisitions were performed by a **DJI Phantom 4 Pro** UAS equipped with RGB CMOS 1" sensor of 21 megapixels resolution, Field of View (FOV) 84°, 8.8 mm/24 mm (35 mm format equivalent), f/2.8-f/11 autofocus from 1 metre to ∞, utilising 12 Ground Control Points (GCPs), to obtain high accuracy orthomosaics and perform the temporal tracking process described in the following. In each acquisition date, an orthophoto of the entire field was generated and labeled starting from images taken by the UAS flying between 15 m and 80 m above ground level and with the camera in nadiral position (perpendicular to the ground).

Data Preprocessing

Once the drone acquisition phase was carried out, the images collected during the flight were merged to form a single geometrically correct (orthorectified) and high-resolution image called **orthomosaic**. The orthomosaics construction was made using OpenDroneMap [1], an application and API for drone image processing capable of constructing an orthomosaics from a group of individual georeferenced images.

Figure 1: Orthomosaic and detection of the entire area of the experiment by YOLOv5 at an altitude of 80 m.

Seven orthomosaics corresponding to UAS flights in the months between September 2021 and December 2021 were generated, and an image dataset was extracted from each of them for the network training. After an initial phase of manual labeling performed using the software VGG Image Annotator (VIA) [2] to obtain the ground truth of the data, the dataset generation was performed by randomly cropping orthomosaics at different heights and applying data-augmentation algorithms (rotation, blurring, saturation, etc.) to the obtained images to produce representative samples. The resulting dataset was then divided into training, validation and test sets and provided as input to the detection networks.

Deep Learning Plant Detection

The **Feature Pyramid Network** and **YOLOv5** [3, 4] used for the detection of artichoke plants are particular type of Single Shot Detector. These types of algorithms work in a single forward pass of the network, locating and classifying objects at the same time. The basic concepts of these networks are the use of a grid that divides the image into cells responsible for detecting objects in that region of the image and the use of priors, predefined boxes responsible for detecting objects of specific sizes and shapes within a grid cell.

COMP4DRONES

ECSEL Joint Undertaking (JU) - agreement No 826610

The main structure of the **FPN architecture** is composed by a bottom-up pathway for the feature extraction and by a top-down pathway for position detection on the image. The combination of these two phases allows the network to detect objects of different scales with a good level of location precision in rapid training times, making this network robust for different drone heights. Our implementation: **Grid Sizes**: (4x4, 8x8, 16x16); **Priors Sizes**: (1x1), (2x1), (1x2); **Input Size**: (512x512); **Total params**: 2.8 M; **Loss**: Boxiness + Location; **IoU**: 50%; **Optimizer**: Adam.

Table 1: Results of the detection with FPN.

Date	Precision	Recall	F_1 score
09_07	0.977	0.868	0.919
09_14	0.960	0.910	0.934
10_01	0.933	0.891	0.912
10_09	0.968	0.905	0.936
11_08	0.935	0.765	0.842
12_03	0.860	0.783	0.820
12_23	0.878	0.843	0.860

As for the **training** of **YOLOv5**, the nano version was trained in 200 epochs and with a batchsize of 8. The results obtained are similar to those reported in Table 1 for the FP network.

Real Time Performance

In the operational phase only the YOLOv5 network (lighter and with less consumption) will be mounted on a NVIDIA® Jetson Nano™ board, embedded in the drone.

Figure 2: YOLOv5 detection performance detail.

The performance in terms of frames per second (for the YOLOv5 network in TRT format) is between 20-24 fps.

Multitemporal Analysis

To correctly reconstruct the complete time history of each plant, a **temporal tracking** algorithm has been developed. It is an automatic registration process, based on the information provided by the box-plants detected by the NNs, searching for the overall translation that maximizes the IoU between the boxes detected in the two consecutive orthomosaics. The result of this tracking process is a series of complete traces, from the first orthomosaic recorded, up to the last available, for all the box-plants that have been detected by the NNs.

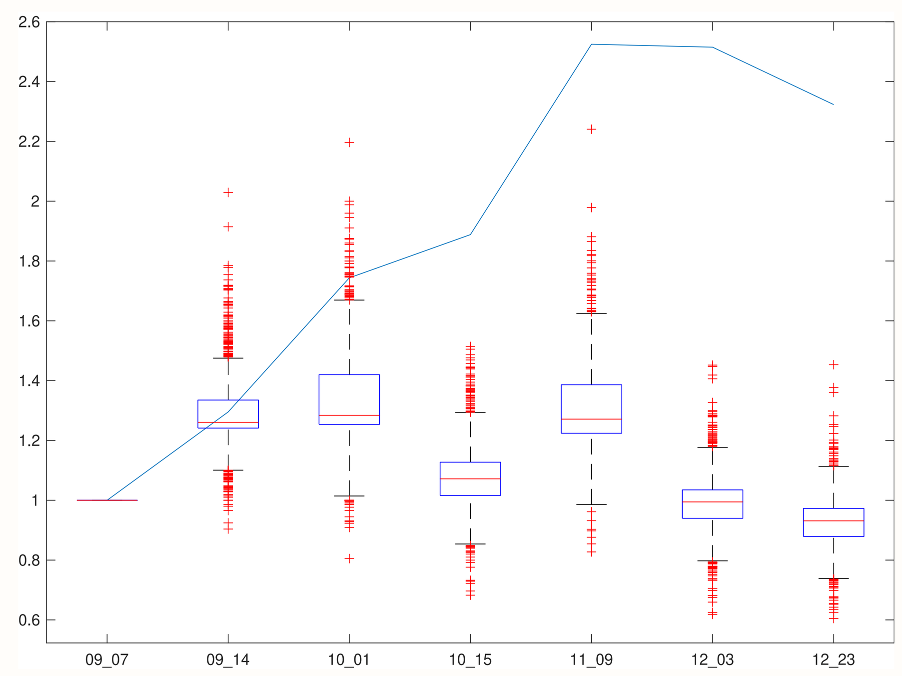

Figure 3: Evolution **growing index**, computed as the ratio between the average size (width and height) of the bounding box detected at the different date.

References

[1] https://github.com/OpenDroneMap/ODM

[2] https://doi.org/10.1145/3343031.3350535

[3] T. Y. Lin et alii, *Feature Pyramid Networks for Object Detection*, 2017 IEEE CCVPR

[4] https://github.com/ultralytics/yolov5

Cyber Physical Systems Summer School 2022, September 19-23 Pula, Sardinia, Italy