Deliverable 5.3: Plan & report of AI weed mapping for AWM (version 2)

Grant Agreement number: 101083589
Start date of the project: 01/05/2023 **End date of the project**: 30/04/2027
Deliverable due date: 31/10/2024 **Date of delivery**: 30/10/2024
Classification: Public
Associated Work Package(s)

WP1	WP2	WP3	WP4	WP5	WP6	WP7	WP8
☐	☐	☐	☐	☑	☐	☐	☐

Version History

Version number	Implemented by	Notes
1.0	UG	
2.0	UG	

Image processing activities and AI pipeline for weed detection ... 2
Introduction: experimental design ... 2
Annual crops ... 2
Perennial crops ... 2
Important weeds and level of discrimination detail of LLs ... 2
Data collection ... 4
Image processing ... 5
Orthomosaic generation ... 5
Image annotation ... 7
Modelling ... 8
Input data ... 8
Model evaluation ... 9
Insights from AI mapping useful to interpret AWM solutions ... 10
Practical approach: situation sketch and results ... 10
Data processing ... 10
Image acquisition ... 10
Image processing ... 10
CVAT: available datasets and annotations ... 11
Preliminary results and insights of deep learning approach ... 12
Supervised learning ... 12
Unsupervised learning ... 15
Conclusions ... 17
References ... 18

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Image processing activities and AI pipeline for weed detection

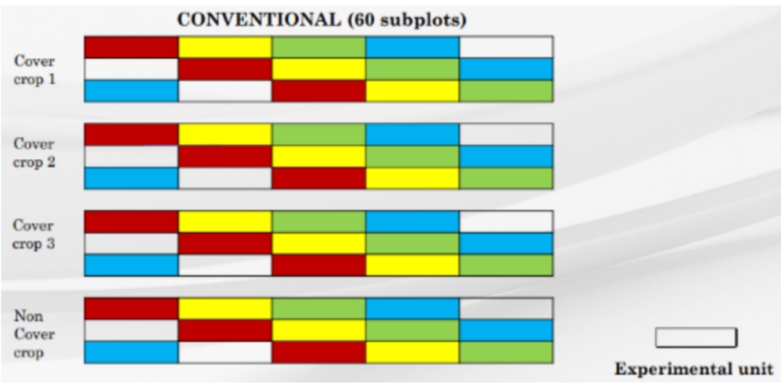
Work package (WP) 5 of the GOOD project deals with digitalization technologies of agroecological weed management systems. Here, we report on the progress on Task 5.1, automated weed identification and mapping. This automated weed mapping can eventually be applied for site-specific weed management. Very high (mm) to high (cm) resolution RGB data are collected with UAS (unmanned aerial system) technology and are then analysed using deep learning, either in a semi-automated (supervised) or automated (unsupervised) way. This results in weed density maps of either all weeds present in the field, of different genera of weeds, or of different species of weeds. The latter aspect depends on the exact requirements for weed control in the individual living labs, which are introduced below. The application is also different in annual and perennial crops.

Introduction: experimental design

Annual crops

For each LL's main crop, three cover crops and one control plot (without cover crop) was established in the first year. This 1st year is a pre-selection stage to select the most suitable cover crop for further experiments in years 2 and 3, with or without AMF inoculation.

The use of cover crops will be combined with other weed control strategies applied to the main crop. For each pilot site, these treatments will be based on the levers described in Section 2, and include, at least, 1) one cultural, 2) one mechanical, 3) one chemical "standard chemical practice"), 4) one chemical at reduced rate, and 5) one control treatment without any weeding.



Perennial crops

Three cover crops (single species or mixtures: legume, grass, cereals, etc.) are sown on the interrows (corridors) of the orchard or vineyard; a reference is kept without cover crop ("interrows standard practice"). Four additional practices are experimented on the interrows: herbicide–full rate; herbicide–half rate; one nonchemical practice (e.g., mulching, mowing, etc.), plus an untreated control (weedy)

There will be 7 treatments in total for the conventional pilot sites with a minimum of 3 replications per treatment (21 subplots).

The treatments for both the conventional and organic & mixed sites are specified in the Living labs calendar.

Important weeds and level of discrimination detail of LLs

A questionnaire was sent to each LL to inquire about the exact purpose of the weed mapping system; feedback was received from all but two LLs. The outcome per LL is presented in **Table 1**. This reveals that for an efficient weed management, in all LLs, weeds should be discriminated to at least the genus level. This level of discrimination determines to a large extent the required resolution of the UAS data,

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

since it is not only sufficient to detect weeds as such, e.g. as plants (green objects) in between rows, but for species/genus level discrimination, the resolution should be sufficient to distinguish between genera, looking at the shapes of the leaves. In row crops, where weed treatment is typically done early after germination, this requires a resolution of at least 5mm. It cannot be guaranteed that this level of detail is feasible in all living labs, given the available UAS systems.

Further, the questioning also confirmed the diversity of the weed species to be discriminated within the project. As expected, there are large differences between the species that should be discriminated according to region and crop type.

Table 1: Results questionnaire of the most damaging weeds and required level of distinction for each living lab

Country	Crop	Code	Most important weed species	Classification detail
Cyprus 1	Olives	CY_olives	/	/
France	Apple-Plum	FR_apple-plum	Broadleaf weeds, monocotyledon weeds	Species, at least dicotyledons vs monocotyledons
Greece1	Wheat	GR_wheat	*Avena*, *Lolium*, *Sinapis* and *Galium*	Species
Greece2	Grapes	GR_grapes	All the perennial weeds	Species
Italy1	Triticale	IT_triticale	*Chrysanthemum coronarium*, *Oxalis cernua*, *Malva* sp.	Genus
Italy2	Citrus	IT_citrus	*Oxalis pes-caprae* L. - *Urtica dioica* L.	Species
Italy3	Grapes	IT_grapes	*Cynodon dactylon* L.	Species
Latvia	Rye-Pea	LV_rye-pea	*Agropyron repens*, *Galium aparine*, *Matricara perforate*, *Poa annua; Bromus secalinus*	Genus
Netherlands	Onion	NL_onion	*Chenopodium album*, *Poa annua*, *Galinsoga* spp.	Species
Portugal1	Cowpea	PT_cowpea	*Cytisus scoparius*, *Cistus ladanifer*, *Lavandula stoechas*	Species
Portugal2	Olives	PT_olives	*Cytisus scoparius*, *Cistus ladanifer; Lavandula stoechas*	Species
Serbia1	Soybean	RS_soybean	*Chenopodium album*, *Sorghum halepense*	Species
Serbia2	Maize	RS_maize	*Chenopodium album*, *Sorghum halepense*	Species
Spain1	Rice	ES_rice	*Echinochloa* spp., *Leptochloa* spp, *Cyperus*, *Ammania*	Species
Spain2	Cherry	ES_cherry	*Lolium*, *Echium*, *Sonchus*, rumex	Genus
Spain3	Apple-Grapes	ES_apple-grapes	/	/

The actual detected weed species will be reported in the Deliverables of WP2-WP3.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Data collection

The image gathering is done with high resolution RGB sensors onboard UAS (unmanned aerial systems). A document with practical guidelines for performing the flights was compiled and was shared with all LLs. The separate LLs were also contacted for further guided support with the organisation of the UAS flights. The goal of the document was to ensure that image collection was done in a standardized way, and that all data were of high quality and readily available for post-processing. These guidelines include the following information:

Flight approach: Mapping flights

It was decided (see further) that we want to work with orthomosaic imagery, a single composite of the individual images covering the entire field, and not with individual UAS images. This requires mapping flights, with sufficient overlap between individual images to create orthomosaics. Mapping flights are best executed automatically making use of the waypoint controls software. This can be done by a preprogrammed flight using the DJI Pilot, DJI Pilot 2 or other dedicated software. It is important to use sufficient overlap between the images, with at least 70% in horizontal and 70% in vertical direction (75% if feasible).

Camera resolution

To ensure the ability to map the weeds, a minimal resolution of 20 MP is needed. For this application, possible drone-sensor combinations are the DJI Mavic 2 series, DJI Mavic 3 series (DJI Mavic 3 Pro; DJI Mavic 3 MS), or Zenmuse P1 gimbal on DJI M300 or DJI M350; other combinations are of course also possible. Note that the resolution is also determined by the flight height (see further), and that the quality of the imagery is not only a function of the resolution of the sensor, but also of its intrinsic qualities (sensor size, lens type, optical quality, data storage, sensor settings, flight conditions, etc.)

Flight height

The discrimination between weeds requires high resolution images. To obtain this, the minimal height possible for a flight should be selected. For a standard mapping flight with DJI software, this is 12m above the ground.

It is important to note that safety always should be the main priority. At low flight heights, special attention needs to be paid to the presence of obstacles (trees, poles, buildings) surrounding the field. If required, either the flight height must be increased to safe levels (at least 3 m above objects that need to be overflown) or the area needs to be adjusted (keeping a safe distance of at least 5m from taller objects).

Ground control points

Ground control points (GCPs) are reference panels put in the field to extract the exact location for the orthomosaic creation. Commonly, a 2x2 black-and-white checkerboards are used, of which the location of the centre is extracted with a RTK-GNSS system. It is recommended to put 6 to 8 GCP's per flight, evenly spread out over the flight area. The use of GCPs is important for improving the image alignment and for later orthomosaicking, particularly in our case of repeated flights.

Flight conditions

If possible, fly close to solar noon (less shadows) or in light overcast (but still bright enough) conditions. Reduce the flight speed if conditions are cloudy, to avoid motion blur. Ideally, keep the shutter speeds small (<1/1000 of a second) and your ISO setting relatively low. In changing weather conditions, set your camera to AUTO-ISO, so that every image is bright enough. Set your camera to -0.5 or -0.7 stops to avoid saturation of brighter plants.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Data are collected by each individual LL following these guidelines.

Image processing

The collected data consist of sets of individual images per executed flight. In GOOD, we decided to work with the orthomosaic rather than the individual images, although this requires an additional processing stage. The reason for this choice is that the orthomosaic option has several advantages: *i*) every single individual plant is imaged in a near-nadir position, offering a more homogeneous dataset; *ii*) the total size of the dataset is much reduced, which is more efficient for data labelling, handling and storage; *iii*) the orthomosaic is precisely georeferenced, and therefore all products generated with the mosaic will also be accurately georeferenced, and *iv*) this orthomosaic can be processed immediately to field maps of weed density and further on to weed prescription maps, as was promised in the GOOD project proposal. Apart from site-specific weed management, georeferenced maps also allow precise monitoring of weed through time. It was decided that UG will generate the georeferenced orthomosaics for all LLs. Starting from this orthomosaic, several processing steps will be performed, to prepare smaller, individual images, suitable for modelling.

Orthomosaic generation

To make the process clear, we will use a specific example (Fig 1a). Using a DJI Mavic 2, a typical flight can give rise to about 400 images, depending on the hight and area of the flight. Each image with this drone is 20MP.

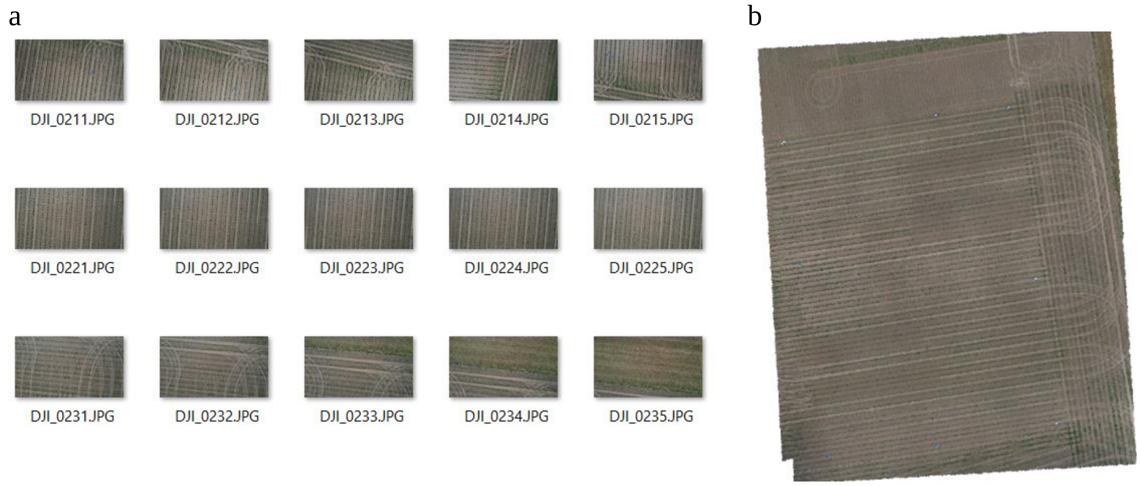


Figure 1: a) Example of several images taken by the DJI Mavic 2 drone; b) the resulting orthomosaic, covering an area of 76 by 80m.

<u>Image mosaicking</u>

The individual images are converted into the orthomosaic and digital elevation model (DEM) using dedicated structure-from-motion software, in this case Agisoft Metashape Professional (Version 1.8.4) will be used. All images have a certain degree of overlap. Correspondence between the images is found using feature detection, commonly scale invariant feature transform (SIFT). Once the features have been detected, they are matched between different images. The first step of the process is the image alignment, resulting in a sparse 3D point cloud and a first estimate of the camera positions. After this, GCPs can be used to georeference this alignment precisely and further correct the camera positions. The next steps include the generation of a dense 3D point cloud, mesh, digital elevation model, and finally a georeferenced orthomosaic that can be exported (Fig. 1b). Unless processed otherwise, this orthomosaic will have the same ground sampling distance (ground resolution resolution) as the original images.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Sampling strategy

An orthomosaic is subsequently processed for annotation and model training. It is not feasible or necessary to annotate the entire field; instead, a subselection of tiles (subimages), each of the same size, is taken. This subselection is done through:

1) The experimental design of the specific pilot field.
2) Input from the LLs weed scientists, with regard to the most important weed patches to annotate.
3) Finding a representative distribution of low to high weed/crop density.

A schematic overview of the sampling process is given in Figure 2.

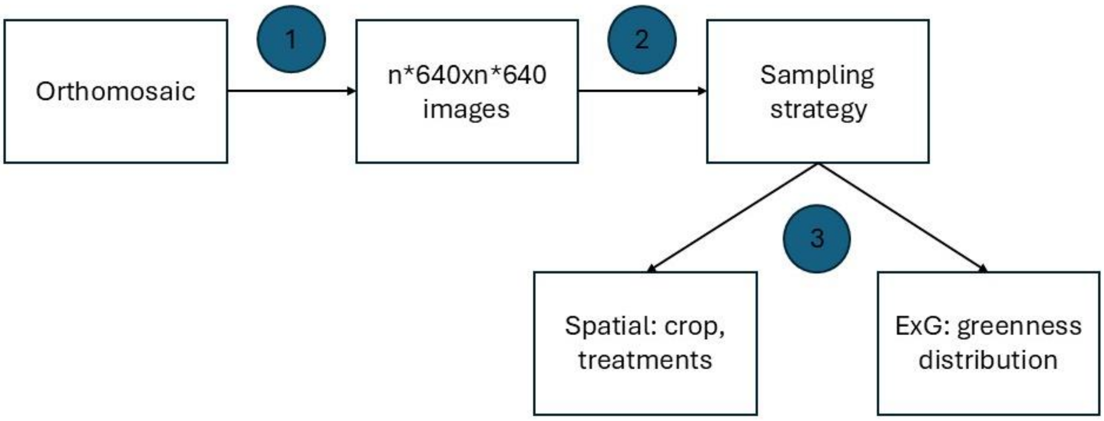


Figure 2: Overview of the sampling strategy. Step 1 involves dividing the original orthomosaic into tiles of size n × 640 by n × 640, determined by the ground sampling distance (GSD). Step 2 illustrates the general approach to sampling these tiles. Step 3 presents two main strategies: (1) dividing tiles based on spatial factors, such as crop type and relevant treatment variations within the field, and (2) selecting tiles based on weed/plant abundance by analysing the greenness distribution.

The subsampling needs to take the experimental units into account. With this information, a stratified random sampling is used to make sure that every experimental unit has at least one, and preferably several, annotated tiles. Also, the weed scientists will give input to which experimental units are more useful to sample annotated tiles from. As an example, experimental units without treatments might show a more diverse and more interesting weed flora in comparison to the application of full herbicide rate. Lastly, an orthomosaic represents a diverse environment where the density of weeds and crops can vary. Since deep learning models need to have a representative input space with regards to the test data, it's important to acknowledge this variability. Weed and crop density can be estimated by analysing the greenness of an image tile, with the Excess Green Index (ExG = 2*Red – Green – Blue) being a useful metric. The sampling method selects images based on the ExG distribution (histogram) across the entire field, ensuring the sampled data maintains the same distribution. This approach allows flexibility in adjusting the weights of specific histogram bins, for example, sampling fewer images from areas with low greenness compared to their presence in the field. In each selected tile, all weed plants need to annotated.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Image annotation

Image annotation (labelling) is important for two reasons. First, it allows developing a supervised deep learning algorithm using these labels. Second, it also allows to evaluate the performance of such models, as well as of semi-supervised or unsupervised models. Manual labelling is labour-intensive; therefore, only a small subset of the actual orthomosaic can be labelled. The amount of available images for annotation is related to the resolution of the orthomosaic. 50% of the tiles, produced by cutting up the orthomosaic are uploaded on the CVAT. Depending on the weed density, LLs will annotate at least 300 tiles. Proper training is required to ensure that all different annotators follow the same guidelines, because consistency in annotations is of great importance for the best performance of the models and, thus, the precision of the weed maps that are going to be produced.

Guidelines

The guidelines for the LLs on the annotation are provided by Eden and are described in Instructions for annotation with CVAT.pdf. Once the orthomosaic of a specific flight is generated, the optimal annotation strategy will be discussed with the LL. This strategy depends on a few factors such as:

1) The resolution of the generated orthomosaic
2) The estimated amount/density of weeds in the field
3) The specific configuration of the plots, involved in the pilot field

The annotation happens with the web-based tool Computer Vision Annotation Tool (CVAT). This tool allows annotating specific regions that the user defines as objects. In Figure 3, an example is given of this process, where the defined objects are weed plants in between a main crop (maize). For this project, bounding boxes are used to define the location of the weeds. It is very important that every weed in an image is annotated, as its performance heavily depends on abundant and correct annotations.

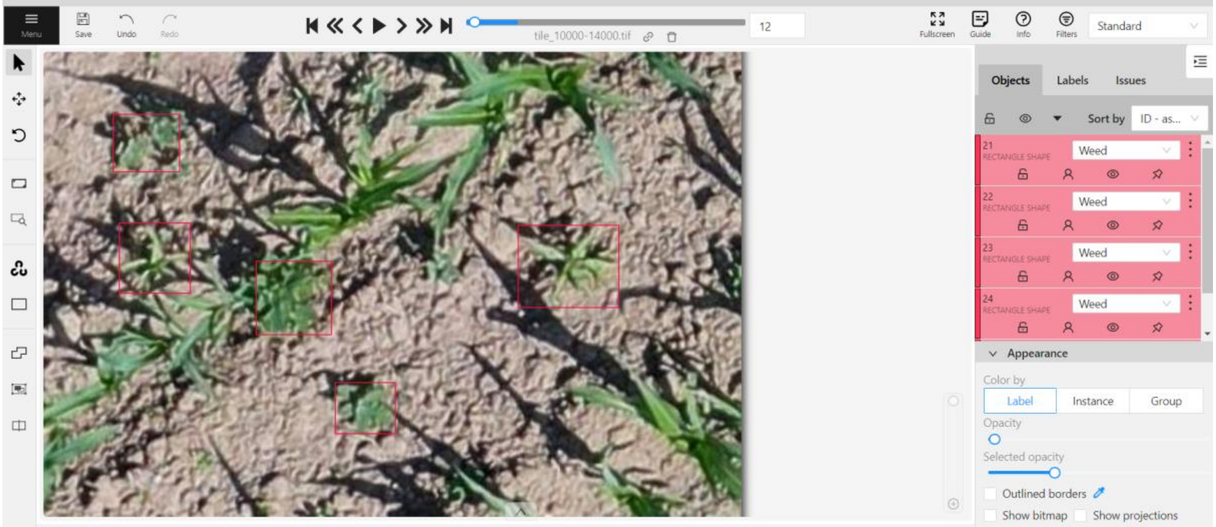


Figure 3: example of the annotation process for a maize dataset.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Annotation format

The annotations, made with the CVAT tool, will be exported in a standardized format. This is set to be the YOLO format. This format stores the annotations as a zip file, which contains all necessary information, such as labelled image names, bounding boxes and corresponding labels.

Modelling

The goal is to map the spatial distribution of the weeds present in the pilot field. This includes the detection, as well as the identification.

Input data

The dataset defined by the stratified random sampling will be used as input for a deep learning model. This consists of the tiles, each with their annotation file.

Supervised model workflow

The general workflow for a supervised neural network is given in Figure 4. The selected tiles through the stratified sampling (dataset) are split into three sets, training data, validation data and test data. The training data is used as input for the model. The model weights are updated by e.g. stochastic gradient descent (SGD), based on the loss function. Parallel to that, validation data is used to evaluate the model, without using this data to update the model. Finally, test data is evaluated by the model after the training is finished. This workflow guarantees that a model does not overfit on the given training dataset. When the model performs well on the test data, it can be used to provide a weed mapping of the unannotated tiles as well as of future datasets of this crop and region.

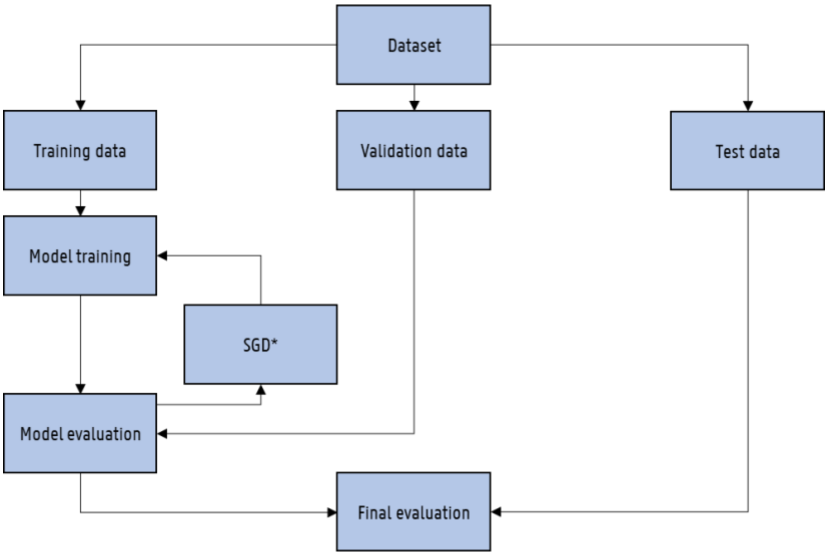


Figure 4: Dataflow of the supervised model.

In first instance, the model of choice is YOLOv8. This is a recent and publicly available deep learning model, with a very powerful architecture. The goal of applying this model is to come up with an accurate deep learning model in little time, so we can already generate weed maps at the entire field level. However, although we will use the same initial architecture for all fields, a specific model will need to be constructed for each LL and crop. The robustness of this model will be verified using consecutive weed datasets, collected later in the season and/or in consecutive years.

Deep learning is evolving very rapidly, and new and more performant models will become available in the near future. We will therefore follow up and test new models in the project. In addition, once we have a large dataset of all the LL and all present weeds, we will explore the possibility to develop one single model, or to develop a model that is pre-trained on a few datasets, but that can then be trained quickly and with relatively fewer labels using transfer learning.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Unsupervised model workflow

Despite the recent advances in deep learning and AI technology, the supervised models mentioned above still rely on a large number of labelled data, and are not able to perform weed mapping on datasets of crops/weed combination that the models were not trained for.

Therefore, UG will also explore an unsupervised model workflow. The idea is that the rows, containing the crops, are detected automatically. With this information, the algorithm can select the plants in-between the rows as the weed plants, and train itself to recognize these plants. For this approach, semantic segmentation was chosen. Bounding box object detection often struggles with predicting densely packed objects. Since we use spatial information in the field, we don't need to compromise by using bounding boxes.

Further, semantic segmentation could allow us to automatically distinguish different clustering groups (species), so that the model can also recognize these different species automatically. These plants can be classified in order to distinguish the different species. The model is then applied to all plants, also the plants close to the rows, so a full understanding of the weed density per class (species) can be provided. This last approach is under research and needs to be confirmed to be achievable.

Model evaluation

The standard evaluation metrics for bounding box object detection are used. For a given bounding box, the model predicts the location and the confidence.

The confidence can be used as a metric to potentially filter out any weak predictions, but this is only based on the model, and not standardly on any ground-truth. The precise location of the bounding boxes will be used as a metric for the accuracy of the model. Model evaluation on the test dataset will be used as a general metric to evaluate the performance. On these data, the intersection over union (IoU), precision and recall will be used as general metrics. The formulae for precision and recall can be seen below, a visual representation of IoU is given in Figure 5.

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

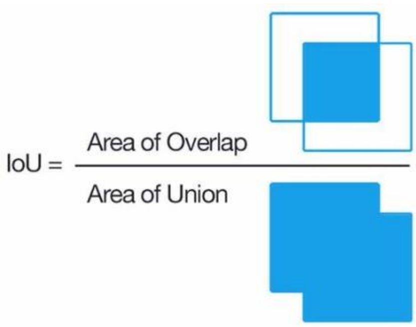


Figure 5: Visual representation of the intersection over union (IoU).

These metrics can be extended to apply on semantic segmentation, meaning they can be used for pixel-level predictions. This allows us to interpret the results of the unsupervised approach. However, the

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

means of metric calculation are completely different, so in order to accurately compare both model types (supervised vs. unsupervised), semantic segmentation masks will be converted to bounding boxes (where possible, in case of extreme weed density, this is not possible).

Insights from AI mapping useful to interpret AWM solutions

Once weed density maps are generated, they will be used throughout the project for several applications:

1) Models as input for site-specific weed management
   The main purpose of the digital maps is that they can eventually be used to automate weed management interventions. This is of course specific for each crop-weed species case, but based on the density maps and the acceptable weed pressure, the field can be divided into blocks, corresponding to the achievable minimal management unit, with each block being assigned the most suitable weed management (e.g., dosage of chemical spraying; necessity to perform weed mechanical intervention or not, …).

2) Understanding of treatment levels (cfr supra: cover crop, weed treatment) on weed density per weed species.
   The maps will give an overview of the weed density per species/genus for each individual treatment block. This information is likely to be more precise than the manual weed plot inventories executed. This provides the researchers with a very good opportunity to evaluate the effect of the different treatment on the weed density. Further, once the method has shown to be operational and trustworthy, it can save a lot of time by reducing the inventories.

3) Evolution (dynamics) of weed growth through time per treatment level
   Unique to the GOOD project is that the experiments will be repeated during three consecutive growing seasons. The data are georeferenced, so we can monitor the evolution and dynamics of the different weed species over time.

Practical approach: situation sketch and results

Data processing

Image acquisition

As described in Data collection, LLs collected data based on the practical drone guidelines. In the following section, only LLs who have sent their UAV data before **01/11/2024** are mentioned.

An overview of the received datasets is given in Drone flights and Annotations. Out of 17 LLs, 13 have sent UAV based data to UGent. Generally, datasets were not always collected following the provided guidelines. The most common problems are:

1) LLs that do not respect the timing of the flights: multiple datasets were gathered during a late to very late stage of crop development. These datasets are consequently not usable when the goal is to detect individual weeds using an object detection model.
2) LLs send UAV based data, but communicate poorly in terms of data contents.
3) Datasets are collected at an inappropriate resolution, due to higher flight height than requested.

Image processing

To clarify the process described in Image processing, the step by step procedure is shown for an example dataset. Given the dataset of Serbia, captured with a DJI Mavic 3 Multispectral, we start by processing the images into an orthomosaic using Agisoft Metashape. In total 1760 RGB images were provided, each with shape 5280x3956. Using Agisoft Metashape, an orthomosaic was produced with ground sampling distance (GSD) of 3.37 mm/pixel (Figure 6).

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

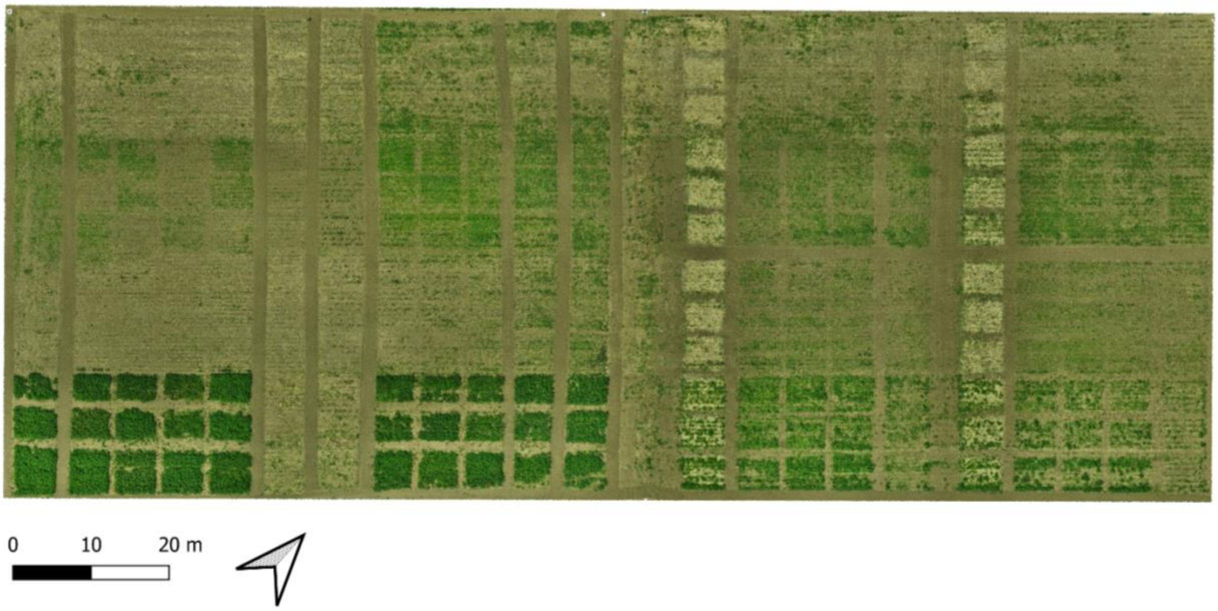


Figure 6: Orthomosaic of The Serbian LL containing Soybean and Maize.

Step 2 is splitting the orthomosaic into n*640 x n*640 tiles. A GSD of 3.37 mm/pixel provides a tile with 2.16m side length when n is chosen to be 1. The orthomosaic was cut into 1161 tiles of 640x640 pixels. In step 3, 50% of these tiles are selected based on the ExG per image, whereafter the images without any vegetation are removed. In this case, this led to a selection of 491 tiles for annotation. In the next step, these tiles were uploaded into CVAT for annotation. The whole pipeline regarding image processing will be done by UGent.

CVAT: available datasets and annotations

Each dataset, processed as mentioned above, is uploaded on CVAT. In Drone flights and Annotations, the date of the upload is also given. Eden is responsible of assigning each of the LL experts for their own LL dataset. A naming convention was also introduced: COUNTRY_CROP_DATE_DRONE:

- **COUNTRY**: A two-letter code representing the country where the data was collected (e.g., "LV" for Latvia).
- **CROP**: The crop type (e.g., Maize).
- **DATE**: The flight date in the format YYMMDD (e.g., "240612" for a flight on June 12, 2024).
- **DRONE**: A unique identifier or code for the UAV used for the data collection (e.g., "M3M" for DJI Mavic 3 Multispectral, "M600" for DJI Matrice 600).

The datasets that have already been uploaded on CVAT and assigned for annotation are the following:

- NL_Onion_240613_Colijnsplaat_zenmuse
- ES_Rice_240410_Conventional_MAVIC2-ENTERPRISE-ADVANCED
- ES_Rice_240410_Ecologico_MAVIC2-ENTERPRISE-ADVANCED
- RS_Maize_conventional_240527_M3M
- RS_Maize_organic_240527_M3M
- RS_Soybean_conventional_240527_M3M
- RS_Soybean_organic_240527_M3M
- LV_CoverCrop_241005
- LV_CoverCrop_241009
- IT_Triticale_240428_M3E
- GR_Wheat_240611_M3M
- GR_Grape_240701_M3M

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

- PT_Olive_240725_Phantop4Pro_Parcel1
- PT_Olive_240725_Phantop4Pro_Parcel2
- PT_Olive_240726_Phantop4Pro_Parcel3
- IT_Grape_240515_M3M_Palizzi_Altomonte
- IT_Grapes_240515_M3M_Palizzi_Santangelo

It is important to note that every dataset will remain uploaded on CVAT for 2 months. Then, Eden will be responsible for deleting them for storage reasons. Also, all partners involved will be responsible for storing them locally (Eden for uploading them on Eden Library repository where they will be openly accessible, UGent and UCD for training AI models and every LL for their own record purposes).

Preliminary results and insights of deep learning approach

Supervised learning

Three preliminary Python scripts were developed for the supervised learning approach for weed detection in the aerial images gathered from the living labs.

The first script is developed to 'tile' the image and annotation datasets from CVAT. Tiling is the process of taking a large image file and splitting it into several smaller size image files. This function is necessary for the training of a YOLOv8 based object detection model to detect weeds in the images. Since YOLOv8 training only processes images up to the size of 640x640 pixels it is necessary to split the larger images to fit this size.

The first script begins by taking a full-size image (1000x1000 pixels in this case) and displaying the hand annotated bounding boxes drawn by the living labs as seen below. The following images were produced by splitting an orthomosaic into 1000x1000 tiles, based on a flight performed with a DJI Mavic 2 on 19/07/2022 by UGent.

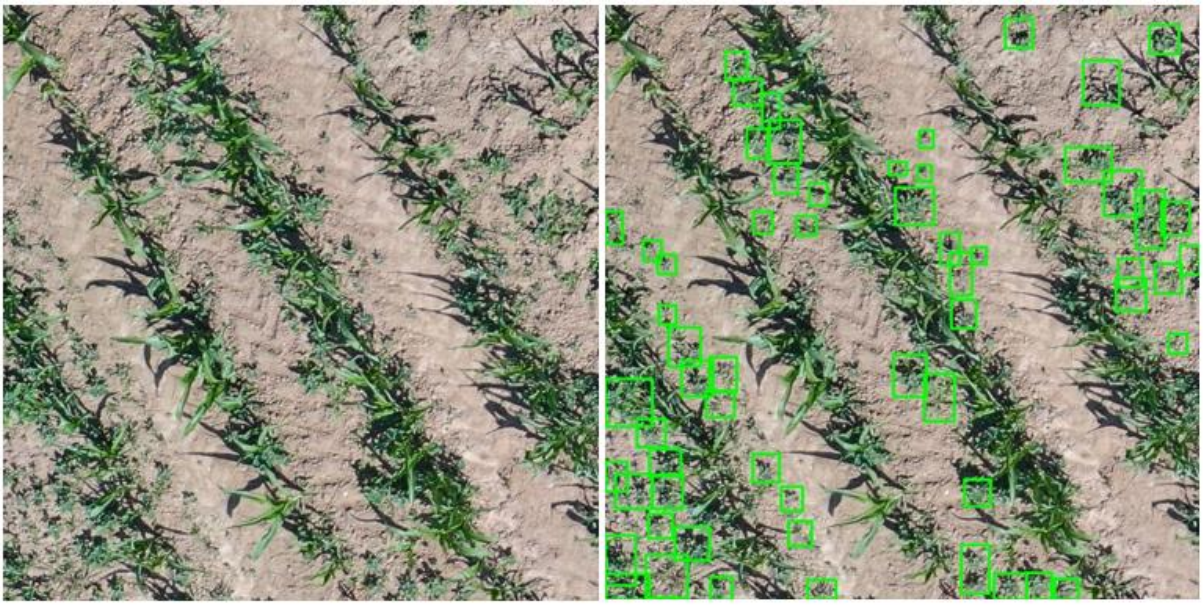


Figure 7: Aerial Image (Left) and Aerial Image with Associated Hand Annotated Bounding Boxes (green) visualized (Right).

The script then takes each image with its associated annotation text file as inputs and splits the image into tiles of size 640x640 pixels. These tiled sections can overlap depending on the size of the input image.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

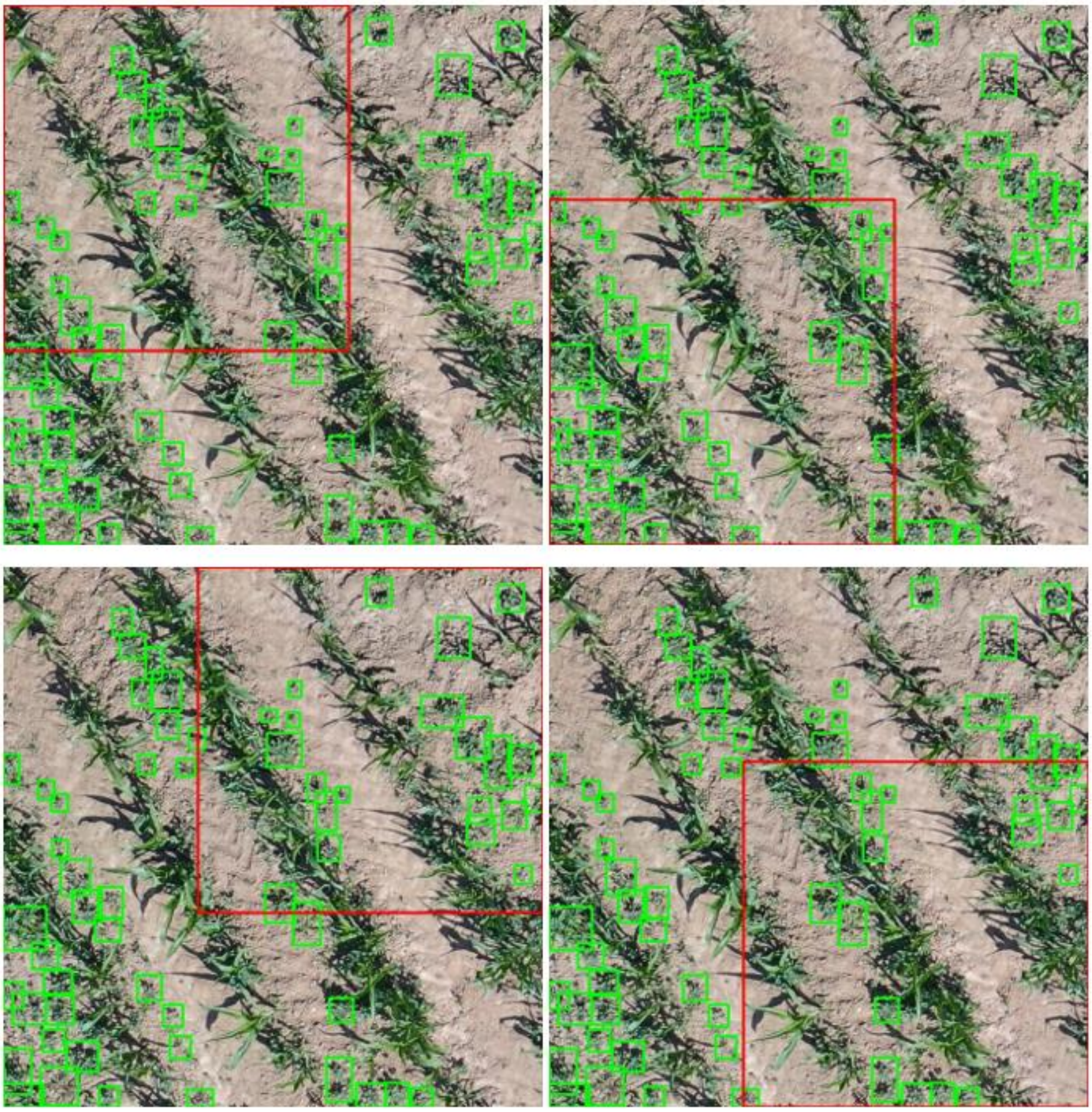


Figure 8: Aerial Image with Associated Hand Annotated Bounding Boxes Visualized with the Tiled Sections Outlined .

The annotations of each 640x640 tile are generated using the annotation text file of the original 1000x1000 image. The annotations are shifted to match the new tiled images or split into separate annotations if a bounding box crosses the different sections boundary lines.

The second script is developed to confirm that the shifted annotation text files have been correctly translated from the original annotation file. The values associated with the annotation boxes are linked to the full image originally and are transformed to be associated with their positioning with regards to the upper left corner of their respective red boxes. The tiled images and shifted annotation files are taken as inputs, and the annotations are projected as YOLO formatted bounding boxes on the tiled images.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

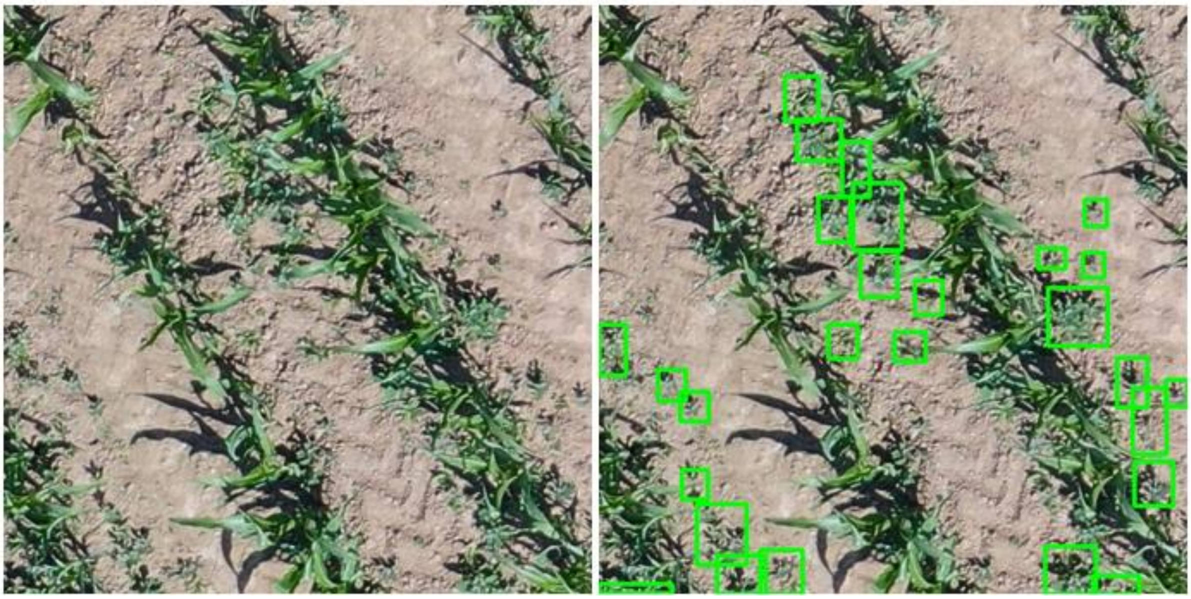


Figure 9: Tiled Image (Left) with Tiled Image and Associated Shifted Bounding Boxes Visualized (Right).

This third script confirms that the annotations text files are properly translated. The tiled images and shifted annotation text files can be used to train a YOLOv8 object detection model. A script was developed to use the image and annotation files for this purpose. This trained model can then be used for object detection of weeds in other sets of full-size images.

A fourth script is currently in development to test the generated object detection models on the full-size aerial images. The quality of the images and annotations being used for the training of the model is what will affect the quality of the predicted bounding boxes for weeds in the aerial images.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Unsupervised learning

Wherever feasible (e.g., row crops, broadleaf weeds in the 2-3 leaf stage), an unsupervised model will be developed to compare against the supervised workflow. Manual annotation is time-consuming and requires specialized agricultural knowledge. Weyler et al. (2024) reported over 800 hours of annotation for a single, fully annotated semantically segmented dataset (across two dates). Unsupervised semantic segmentation of UAV imagery for site-specific weed management has been explored in various scenarios with different sensors. The main advantage of this approach for the GOOD project is the potential to apply this method to regardless which field and crop, in addition of course to the reduction in annotation time, while still generating accurate weed density maps.

UGent has developed and tested such an approach for multiple datasets, that were collected outside the GOOD project. A specific case study for site-specific weed management was conducted in the summer of 2024, where a maize field in Geerbos, Belgium, was monitored. The focus was on testing the unsupervised deep learning method to automatically segment the field into three categories: background, weeds, and maize.

The flight was performed when the maize was in the 2-leaf stage, on June 12, 2024. The weed density was visually inspected and was estimated to be very low to medium. To effectively map weeds, very high resolution RGB imagery was captured using a DJI Mavic3 Multispectral (RGB: 20 MP). Mapping was performed at 12m above ground level (AGL). After processing, the flight using the Mavic 3 resulted in an orthomosaic with ground sampling distance (GSD) of 3.32 mm/pixel. Figure 10 shows the full orthomosaic.

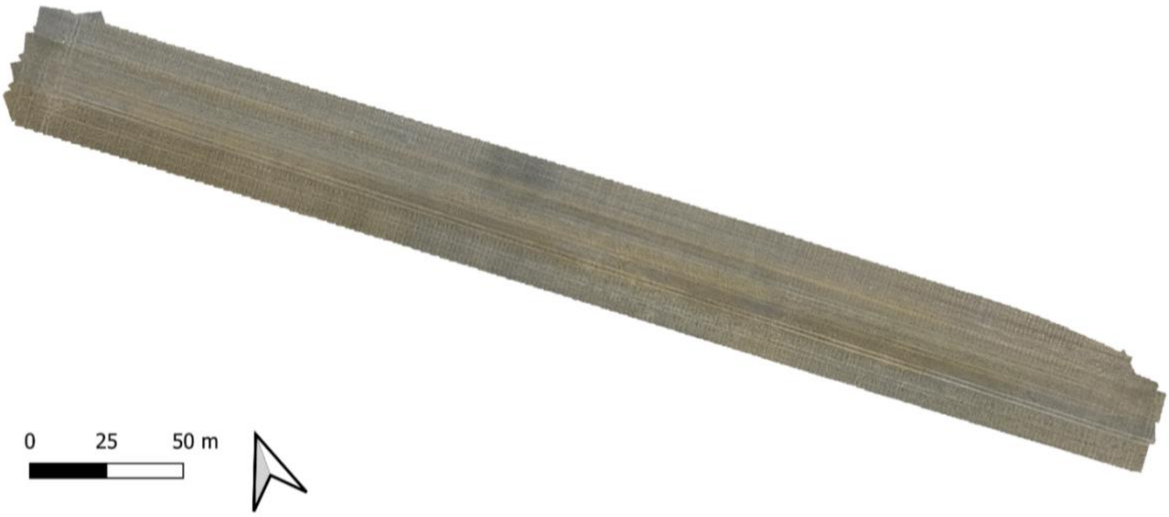


Figure 10: RGB orthomosaic, captured on 12/06/2024.

This algorithm consists of the following steps:

- Semantic segmentation of the ortomosaic in an unsupervised way (no manual labelling), based on vegetation indices.
- Splitting the orthomosaic into subimages.
- Row structure detection within each subimage.
- Creation of foundational segmentation, based on the row structure, whereby the interrow vegetation is classified as the crop of interest, in this case, maize, while the intrarow vegetation is identified as weeds.
- Training of deep learning model using the segmentation information to create pseudo-labels.
- Application of deep learning to entire orthomosaic.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Figure 11 gives a close-up and the resulting output of the model.

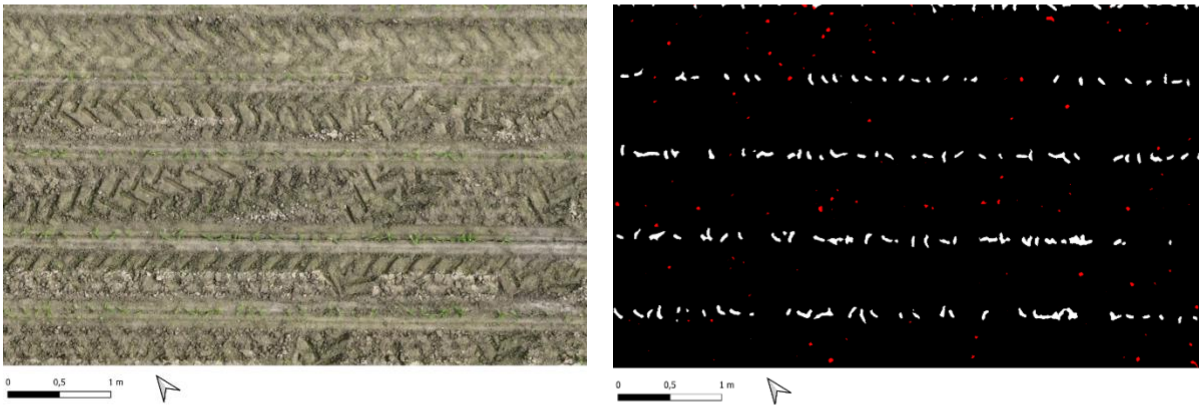


Figure 11: RGB image and the resulting classification of a zoom of the orthomosaic shown in Figure 7

After applying the model on the whole field, the weed density was calculated by dividing the amount of weed pixels by the total amount of pixels, per area. In this case, a 1x1m grid was constructed in QGIS, where the weed density was calculated.

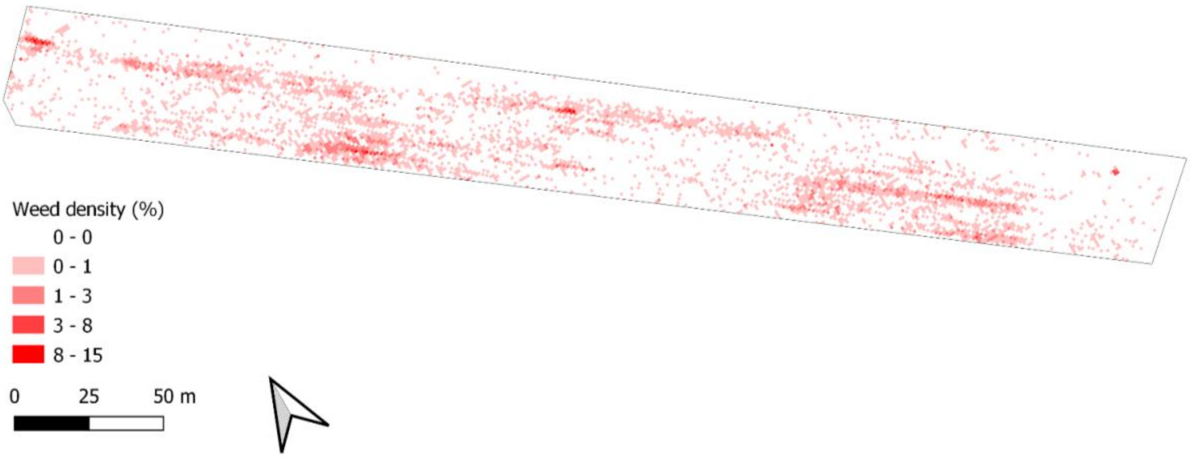


Figure 12: Weed density estimation of the field on 12/06/2024.

In Figure 12, the estimated weed density is presented, categorized into five classes ranging from 0% to 15% weeds. As expected, the weeds are concentrated, in this case in a few stips running east-west. Approximately 38% of all 1x1m blocks have no weeds, while about 60% of the blocks contain less than 1% weeds. For comparison, the average maize density across all blocks is 1.9% ≈ 190 cm^2/m^2. In 61% of the field, the weed density is lower than the maize density.

This method has now been tested on a range of different datasets of different crops, with satisfying results.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

Conclusions

The data gathering, processing, and modelling efforts in the first growing season of the GOOD project encountered several challenges. These barriers were identified during WP5 internal meetings and throughout the data collection process. They were communicated to relevant partners, along with proposed mitigation measures, to ensure more robust data collection in future project years. The most important problems are situated within the data gathering, which is the most critical step in the whole processing pipeline. The most prominent issues lie within the timing, resolution and contents of the data. A plan will be set up to handle these issues in the second growth season of the project. This plan will include general meetings with all LLs to discuss and propose solutions to these types of problems, but will also include LL specific feedback.

The following six months of Task 5.1 will include the annotation of the uploaded images datasets on CVAT, which will be reported in the next Task 5.1 update. These annotations will be used to train LL specific deep learning models, which will also be included in the next update.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.

References

Weyler, J., Magistri, F., Marks, E., Chong, Y. L., Sodano, M., Roggiolani, G., Chebrolu, N., Stachniss, C., & Behley, J. (2024). PhenoBench: A Large Dataset and Benchmarks for Semantic Image Interpretation in the Agricultural Domain. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Funded by the European Union

Funded by the European Union under Grant Agreement No. 101083589. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or REA. Neither the European Union nor the granting authority can be held responsible for them.