



OPEN

DATA DESCRIPTOR

# Manually annotated and curated Dataset of diverse Weed Species in Maize and Sorghum for Computer Vision

Nikita Genze<sup>1,2</sup>, Wouter K. Vahl<sup>3</sup>, Jennifer Groth<sup>3</sup>, Maximilian Wirth<sup>1,2</sup>, Michael Grieb<sup>4</sup> & Dominik G. Grimm<sup>1,2,5</sup>✉

Sustainable weed management strategies are critical to feeding the world's population while preserving ecosystems and biodiversity. Therefore, site-specific weed control strategies based on automation are needed to reduce the additional time and effort required for weeding. Machine vision-based methods appear to be a promising approach for weed detection, but require high quality data on the species in a specific agricultural area. Here we present a dataset, the Moving Fields Weed Dataset (MFWD), which captures the growth of 28 weed species commonly found in sorghum and maize fields in Germany. A total of 94,321 images were acquired in a fully automated, high-throughput phenotyping facility to track over 5,000 individual plants at high spatial and temporal resolution. A rich set of manually curated ground truth information is also provided, which can be used not only for plant species classification, object detection and instance segmentation tasks, but also for multiple object tracking.

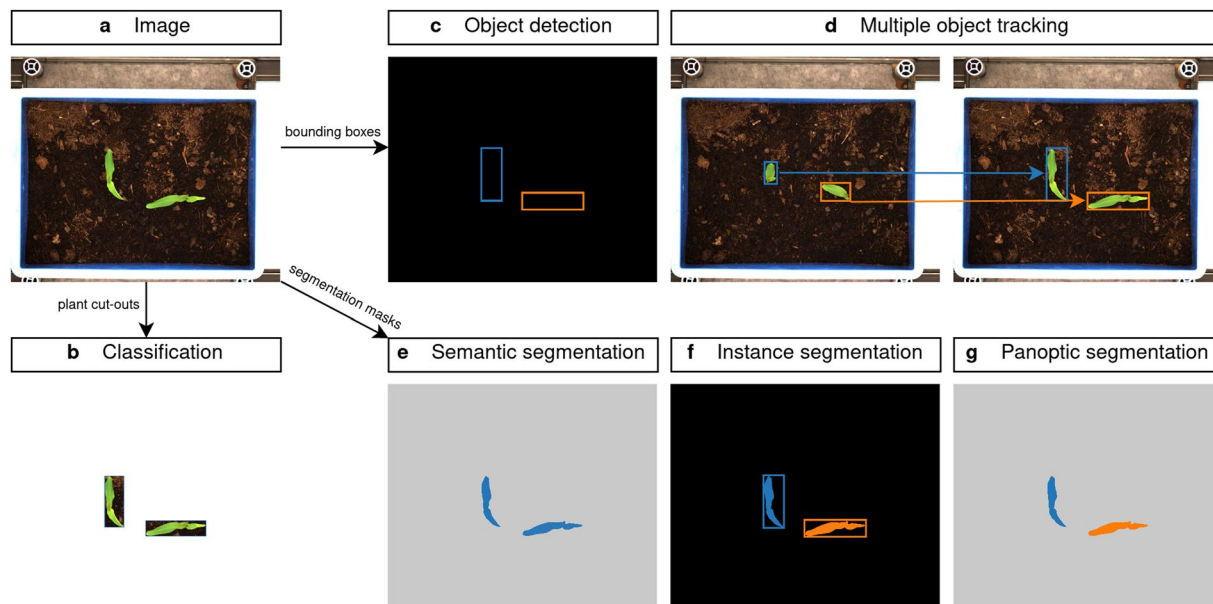
## Background & Summary

Weeds are plants that, although not specifically cultivated, are adapted to grow on arable land. Typically, weeds are considered to be an undesirable element in crop production. Their negative impact on crop development can be described in terms of competition with the crop for resources (nutrients, sunlight, space and water), reduction in productivity, increased challenges during harvesting and an overall increase in the cost of agricultural production. In addition, weeds can be hosts for insects and diseases<sup>1,2</sup>, which might further increase the necessity for control strategies. Nevertheless, weeds might also have positive effects on biodiversity<sup>3</sup> and soil structure<sup>4</sup>. Therefore, only highly competitive and invasive weed species should be removed which might lead to more sustainable agriculture<sup>5</sup>.

Over centuries, several crop management strategies were established to mitigate the negative impact of weeds, which can be divided into five main categories<sup>6</sup>: 'preventative' (preventing weeds from establishing), 'cultural' (maintaining field hygiene with low weed seed bank), 'mechanical' (removing weeds by mowing, mulching or tilling), 'biological' (using natural enemies such as insects or animals), and 'chemical' (applying herbicides). Disadvantages of these approaches include financial burden, additional time and effort to varying degrees. In addition, control treatments may impact the health of people, plants, soil, animals, and the environment<sup>7-9</sup>.

Sustainable strategies for managing weeds are critical to feed the world's population while conserving the ecosystems and biodiversity<sup>9</sup>. The limited and rational use of herbicides is an important principle of sustainable farming, as spraying of herbicides leads to waste and can pollute soil and water sources. Furthermore, agrochemical residues are one of the most important food-related concerns. Therefore, additional non-chemical and site-specific weed management (SSWM) strategies<sup>10</sup> are needed, which should be linked with the farm

<sup>1</sup>Technical University of Munich, TUM Campus Straubing for Biotechnology and Sustainability, Bioinformatics, Schulgasse 22, 94315, Straubing, Germany. <sup>2</sup>Weihenstephan-Triesdorf University of Applied Sciences, Bioinformatics, Petersgasse 18, 94315, Straubing, Germany. <sup>3</sup>Institute for Crop Science and Plant Breeding, Bavarian State Research Center for Agriculture, Am Gereuth 6, 85354, Freising, Germany. <sup>4</sup>Technology and Support Centre in the Centre of Excellence for Renewable Resources (TFZ), Schulgasse 18, 94315, Straubing, Germany. <sup>5</sup>Technical University of Munich, TUM School of Computation, Information and Technology (CIT), Boltzmannstr. 3, 85748, Garching, Germany. ✉e-mail: [dominik.grimm@hswt.de](mailto:dominik.grimm@hswt.de)



**Fig. 1** Comparison of different computer vision tasks used in plant phenotyping. Plants (a) can be classified (b) as individual cut-outs, detected by bounding boxes (c), tracked through time (d), or segmented containing pixel-wise information. This can be done either semantically (e), by instance (f) or by the combination of both called panoptic segmentation (g).

management system. One key aspect is to automatically and precisely detect weeds to mitigate the additional time and effort for either site-specific or weed-specific herbicide application or mechanical weed control. Numerous studies demonstrated methods to automatically detect weeds on the field or in greenhouses, where computer vision-based methods seem the most promising<sup>11–19</sup>. These methods can be grouped into different tasks with varying ground truth information, as shown in Fig. 1. Starting from an image as shown in Fig. 1a, image classification is the least accurate task (Fig. 1b). It is applied to a single plant cut-out without location and thus cannot be applied in SSWM tasks, where the simultaneous detection and localization of multiple plants is desired. Therefore, the object detection<sup>20</sup> task might be utilized. However, this task detects rectangular bounding boxes (Fig. 1c) which is not satisfactory due to the complex and irregular shape of the plants. Also, these models are prone to occlusion<sup>21</sup> which diminishes their performance in areas with high weed infestation<sup>22</sup>. Nevertheless, the analysis of a plant's growth dynamics can be achieved, where multiple objects are tracked through time (see Fig. 1d). Moreover, by using segmentation masks, additional tasks can be performed. Here, by convention any countable entity (i.e. plant, person, etc.) is named ‘thing’ and any uncountable region (i.e. soil, sky) is called ‘stuff’: The semantic segmentation task<sup>23</sup> provides a precise delineation of *stuff*, which separates every pixel in the image by class label, but cannot separate different plants of the same class<sup>24</sup> (see Fig. 1e). This is crucial in selective weed management<sup>25</sup> to conserve biodiversity by removing only competitive weeds. Consequently, instance segmentation<sup>24</sup> (Fig. 1f) can generate accurate detections of *things* individually, which can be used in many downstream tasks such as weed density assessment or biomass estimation<sup>26</sup>. Nevertheless, this type of model is only able to detect countable objects (*things*) and does not consider *stuff* regions. Finally, panoptic segmentation<sup>27</sup> combines the concept of both semantic and instance segmentation and assigns two labels (semantic label and instance id) to each pixel in an image (Fig. 1g).

The basis for the development, validation and assessment of such systems is the availability of high quality data on weed diversity in a particular area of interest. Several datasets are publicly available, but lack several aspects for precise plant phenotyping, as summarized in Table 1.

Most datasets available lack variability in the data (low number of individuals or plant species) limiting their usability in different studies. Only a few datasets are larger than 100,000 annotated plant samples, including Open Plant Phenotype Database (OPPD)<sup>28</sup> and Pl@ntNet-300k<sup>29</sup>. However, Pl@ntNet-300k can only be used for classification tasks without tracking plant growth stages and OPPD is missing semantic and instance segmentation masks, which are important for precise phenotyping. Also, the bounding box information of a plant is often not sufficient, as it is too coarse for most weed management applications. Therefore, semantic segmentation or even more accurate instance segmentation masks are required. Finally, tracking a plant over time provides valuable insight into growth dynamics.

In this work we have created a high-quality dataset of different plant species with a high temporal and spatial resolution. We added manually curated semantic and instance segmentation masks of a subset to make this dataset suitable for weed management tasks. For this purpose, we used a high throughput phenotyping system to ensure a high degree of automation, as this system was equipped with controlled illumination and an automatic irrigation system. In our dataset, we included images of plants captured multiple times per day. This included captures in the evening, when the appearance of some species changes due to their dependency on sunlight. In

Dataset name	Labeled samples	Classification	Object detection	Semantic segmentation	Instance segmentation	Multiple object tracking (growth)
Open Plant Phenotype Database (OPPD) <sup>28</sup>	315,038	✓	✓			✓
Pl@ntNet-300k <sup>29</sup>	306,146	✓				
<b>MFWD (ours)</b> <sup>30</sup>	<b>200,148</b>	✓	✓	✓	✓	✓
Plant Village Dataset <sup>41</sup>	54,303	✓				
Deep Seedling <sup>42</sup>	33,444	✓	✓			
DeepWeeds <sup>43</sup>	17,509	✓				
Weed25 <sup>44</sup>	14,023	✓	✓			
WE3DS <sup>45</sup>	11,544	✓	✓	✓		
Sudars <i>et al.</i> <sup>46</sup>	7,853	✓	✓			
Plant Seedling <sup>47</sup>	5,539	✓	✓			✓
Champ <i>et al.</i> <sup>24</sup>	2,489	✓	✓	✓	✓	
Ladybird Cobbitty 2017 Brassica Dataset <sup>48</sup>	2,245	✓	✓	✓		
CWFID <sup>49</sup>	494	✓	✓			
Leaf Segmentation and Counting Challenge <sup>50</sup>	284		✓	✓	✓	✓

**Table 1.** Comparison of ground-based image datasets for plant detection.

addition, we generated data from different varieties of sorghum and maize, focusing on a wide range of seedling weeds that are also common in agricultural sites where these crops are grown.

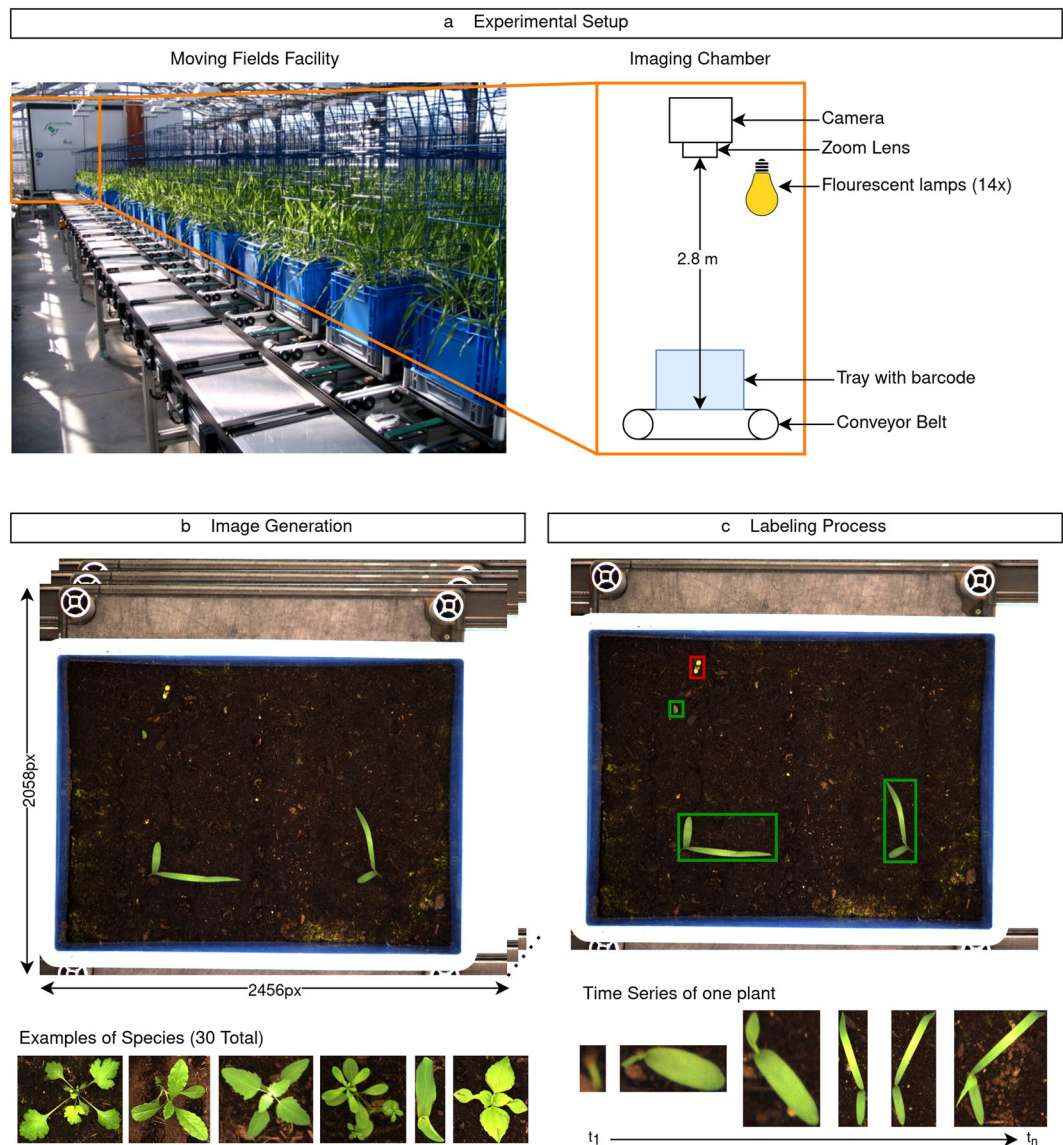
## Methods

The methodology can be summarized into three steps, as shown in Fig. 2. First, we will describe the experimental setup (Fig. 2a). Second, we will illustrate the image generation (Fig. 2b) and conclude with the labeling process (Fig. 2c).

**Experimental Setup.** To generate a dataset consisting of high-quality images that capture the initial growth dynamics of individual plants of several weed species, a greenhouse experiment was performed at the Moving Fields facility (<https://www.lfl.bayern.de/verschiedenes/ueberuns/272457/index.php>) of the Bavarian State Research Center for Agriculture in Freising, Germany. In the experiment, plants were grown in micro-plots that were watered and photographed automatically on at least a daily basis from the day of sowing until harvest, which took place at around shooting. Built into a greenhouse, the Moving Fields facility (LemnaTec GmbH) consists of a conveyor belt system, three irrigation stations and four ‘Scanalyzer 3d’ photo cabins, which together enable experimental units consisting of plants growing on micro-plots to be automatically moved, watered, weighted and photographed. The greenhouse can be climatically controlled with regard to humidity and temperature and can be illuminated by 48 sodium-vapor lamps (Philips Son-T AGRO). The conveyor belt system (Bosch Rexroth TS2plus) accommodates and enables the movement of 390 carriers (micro-plots). At the three measuring stations, digital scales (Bizerba ST) and high-pressure pumps (Wartson-Marlow) enable carriers, together with any plants transported by them, to be weighed and to be watered to a unit-specific target weight. The plant species included in the dataset, listed in Table 2 and Table 3, were selected as weed species common to fields of sorghum grown in Germany. Additional selection criteria were 1) commercial availability and 2) the ability to be grown at the climatically controlled conditions of a greenhouse. Seeds of the species involved were acquired from commercial breeders in Germany, the Netherlands and France. All plants were grown in boxes of size 40 × 30 × 22 cm (outer dimensions). The color of these boxes was blue, to facilitate image analysis afterwards. Each box was filled to about half height (roughly 11 cm after compression) with a commercial peat-free substrate (Höfter GmbH), primarily consisting of coconut fibers. Plants were grown as monocultures; each box contained plants of one species only. To yield data for enough individual plants per species, the number of boxes varied between species due to different germination rates.

The number of seeds planted in each box was made dependent on the expected germination rate, which was adjusted throughout the experiment. Thus, seed density varied both within species over time and between species. Following breeder recommendations, some seeds were kept in a vernalization room (at 4 °C) or treated with gibberellin acid (GA<sub>3</sub>) to ensure germination success. Units were sent to an automatic watering station as often as frequent imaging allowed, in practice at least twice a day. Each unit was watered to its unique target weight. This target weight initially corresponded to the unit’s weight at sowing. Throughout the experiment, target weights were adjusted repeatedly to prevent boxes from becoming either too dry or too wet. Twice a week, all boxes were examined to score seedling emergence, to thin the standing stock in order to minimize overlap between individual plants and to harvest plants that either started shooting or that became too big for the box they were growing in.

In addition, different varieties of maize (*Zea Mays*) and sorghum (*Sorghum bicolor*) were grown and captured, as shown in Table 3.



**Fig. 2** Outline of the experimental workflow. **(a)** Experimental setup using an automated phenotyping facility. **(b)** Generation of images with high temporal and spatial resolution of 30 different plant species. **(c)** Labeling process illustrated by bounding box information for the object detection task (weed illustrated in red and sorghum plants in green).

**Image Generation.** To generate well illuminated, high-resolution top-down images of the experimental units, one of the Scanalyzer 3D imaging cabins of the Moving Fields facility was used. In this cabin, one RGB camera (Basler piA2400-17gm) is mounted 2.8 m perpendicular above the conveyor band. This camera takes images with  $2456 \times 2058$  pixels. This camera is equipped with a motorized zoom lens (Pentax C6Z1218M3-5); throughout the experiment, however, this lens was set fixed at a single position, which resulted in a ground resolution of  $\sim 0.17$  mm per pixel. The micro-plots were illuminated by 14 fluorescent lamps (Osram HE 28 W/865) that were also mounted perpendicular above the conveyor band. Units were imaged as often as possible, in practice at least once a day on the two days per week on which plant maintenance took place and at least twice a day on all other days. The units were tracked over a development period from sowing till the last plant either started shooting or became too big for the setup, representing the relevant stages for weed control on the field for sorghum. The images collected were stored in a LemnaTec-specific raw format, after which they were converted to PNG format. Each experimental unit was marked with a unique numerical barcode composed of identity codes for 1) the species, 2) the treatment (with or without GA3) and 3) the replicate at issue. Each image was saved with the barcode of the unit that was on the image as well as the date and time of image acquisition.

**Labeling Process.** The complete dataset was labeled using the open-source software CVAT (Computer Vision Annotation Tool; <https://www.cvat.ai/>) as a self-hosted solution. This software made the labeling time-efficient, as it provided an easy interface for multiple object tracking by adding time series information



EPPO Code	Dicot	Botanical Name	# Images	Captured Trays	Labeled Trays for		
					classification	detection	segmentation
ACHMI	✓	<i>Achillea millefolium</i>	2,893	25	25	5	0
AGRRE		<i>Elymus repens</i>	2,848	29	12	12	0
ALOMY		<i>Alopecurus myosuroides</i>	3,595	28	11	11	0
ARTVU	✓	<i>Artemisia vulgaris</i>	2,881	34	34	18	1
CHEAL	✓	<i>Chenopodium album</i>	1,652	23	23	9	1
CIRAR	✓	<i>Cirsium arvense</i>	2,546	15	15	15	1
CONAR	✓	<i>Convolvulus arvensis</i>	4,914	27	27	27	1
ECHCG		<i>Echinochloa crus-galli</i>	2,622	27	25	25	0
GALAP	✓	<i>Galium aparine</i>	1,513	23	23	14	1
GASPA	✓	<i>Galinsoga parviflora</i>	847	13	13	13	1
LAMAL	✓	<i>Lamium album</i>	5,276	36	36	32	0
MATCH	✓	<i>Matricaria chamomilla</i>	1,681	21	21	8	0
PLAMA	✓	<i>Plantago major</i>	7,856	43	43	37	0
POAAN		<i>Poa annua</i>	2,574	23	11	11	0
POLCO	✓	<i>Fallopia convolvulus</i>	330	9	9	9	1
POROL	✓	<i>Portulaca oleracea</i>	4,435	21	21	5	1
PULDY	✓	<i>Pulicaria dysenterica</i>	1,745	21	21	10	1
SOLNI	✓	<i>Solanum nigrum</i>	1,472	22	22	12	1
SSYOF	✓	<i>Sisymbrium officinale</i>	2,477	14	14	10	0
STEME	✓	<i>Stellaria media</i>	1,887	22	22	8	0
THLAR	✓	<i>Thlaspi arvense</i>	4,044	38	38	36	1
VEROF	✓	<i>Veronica officinalis</i>	3,987	22	22	5	0
VIORAR	✓	<i>Viola arvensis</i>	1,261	17	17	11	1

**Table 2.** Selected weed species with corresponding amount of data captured and labeled.

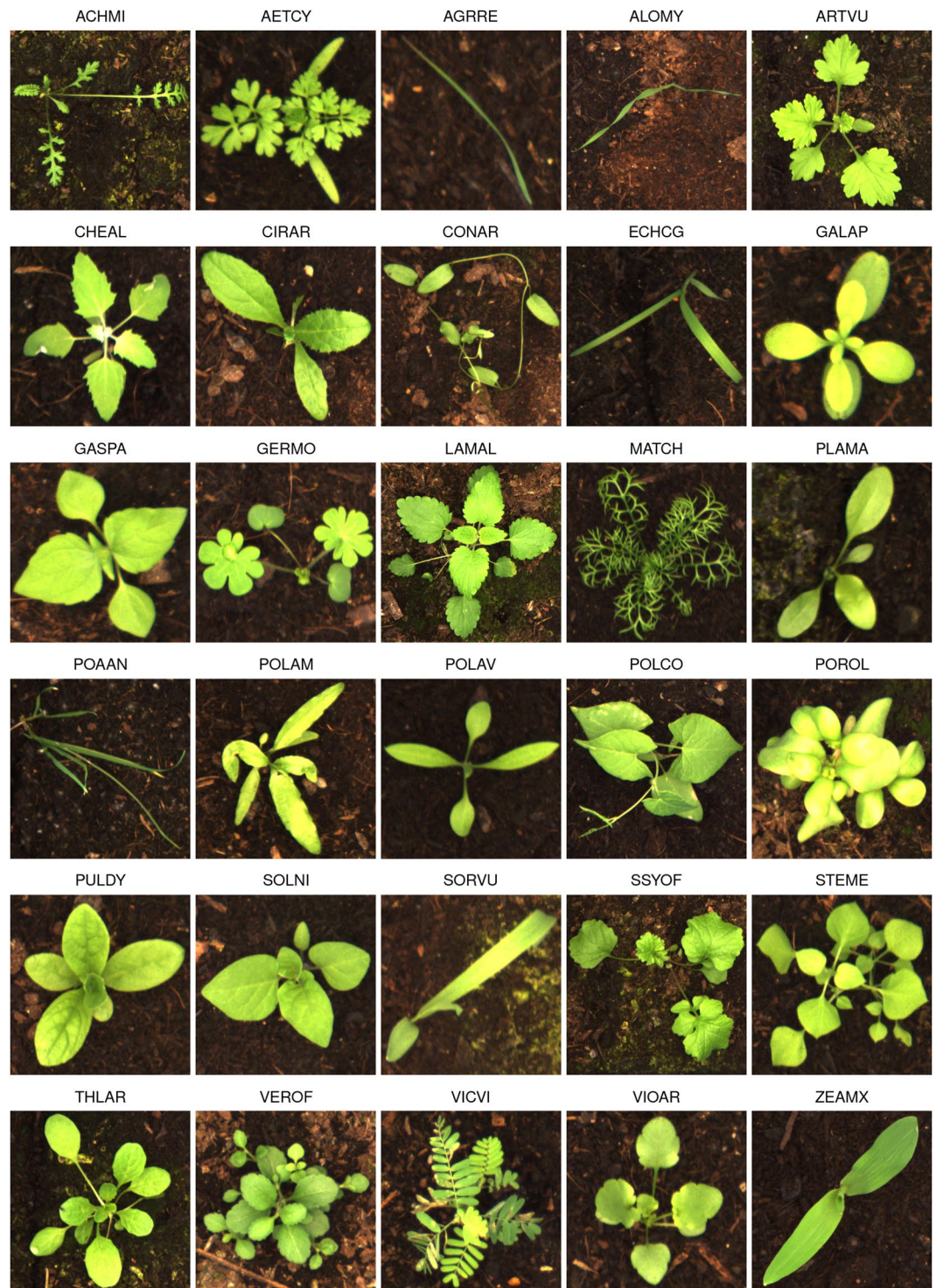
Abbreviation	EPPO Code	Variety	# Images	Captured Trays	Labeled Trays for		
					classification	detection	segmentation
SORFR	SORVU	KWS Freya	2,881	27	27	27	0
SORHA	SORVU	KWS Hannibal	552	27	5	5	1
SORKM	SORVU	KWS Merlin	511	27	5	5	0
SORSA	SORVU	KWS Sammos	1,279	27	12	12	0
SORKS	SORVU	KWS Sole	604	27	5	5	0
SORRS	SORVU	RAGT Swingg	447	27	5	5	0
ZEALP	ZEAMX	Lidea Palladium	930	38	38	38	0
ZEAKJ	ZEAMX	KWS Johaninio	1,660	38	38	38	1

**Table 3.** Selected crop varieties with corresponding amount of data captured and labeled.

EPPO code	Botanical name	# Plant individuals	# Plant samples
AETCY	<i>Aethusa cynapium</i>	4	110
GERMO	<i>Geranium molle</i>	4	139
POLAM	<i>Persicaria amphibia</i>	55	2237
POLAV	<i>Polygonum aviculare</i>	2	62
VICVI	<i>Vicia villosa</i>	1	32

**Table 4.** Additional germinating weed species (dicots) that were not sown explicitly.

per plant. Each species was labeled individually, providing the EPPO code as a label. Although only one species was seeded per tray, more weed species germinated during the experiment (compare Table 4), as the seed assortment was not completely pure. Therefore, the additional label “weed” was used to annotate these plants and the unknown species was identified by an expert in a second step. The correct species could not be specified for all plants due to several reasons (i.e., little germination resulting in small plants, occlusion with other plants, etc.), especially when they were not part of our initial assortment.



**Fig. 3** Example of each plant species with corresponding EPPO code.

The instance segmentation masks were manually drawn using another open-source software, GIMP (GNU Image Manipulation Program), which provides pixel-level information. Therefore, we selected one tray from each of 14 plant species, as not all species could be labeled due to complexity and time constraints (see Fig. 1g).

### Data Records

The Moving Fields Weed Dataset<sup>30</sup> (MFWD) is deposited at the digital library of the Technical University of Munich (<https://mediatum.ub.tum.de/1717366>). The dataset consists of 94,321 high temporal and spatial resolution images of 30 different plant species (see Fig. 3).

Data Content	Data Statistics
High Resolution Tray Images	94,321
Individual Plant Samples (Object Detection)	200,148
Individual Plant Samples (Instance Segmentation)	2,025
Time Series of individual Plants	5,068
Weed Species	28
Sorghum Varieties	6
Maize Varieties	2
All PNG images [GB]	814
All JPEG images [GB]	114

**Table 5.** Summary of the current version of our dataset.

Column	Explanation
track_id	ID of the individual time series connecting bbox_id of different images in one tray
label_id	ID of the plant species, saved as EPPO code
bbox_id	ID of the individual sample
xmin	left coordinate of the individual sample
ymin	top coordinate of the individual sample
xmax	right coordinate of the individual sample
ymax	bottom coordinate of the individual sample
filename	relative file path to the image
tray_id	ID of the tray

**Table 6.** Description of the csv file containing ground truth information.

Additional ground truth data is provided, consisting of the plant species, a bounding box per plant, and time series information to track the same plant individual through growth. We labeled a large subset of these images, resulting in 200,148 records for 5,068 plants in the current version. Additional information is shown in Table 5.

Image data is stored in PNG format to ensure the highest possible quality without compression artifacts. Compressed (JPEG) images are also stored to ensure accessibility with lower Internet bandwidth. All object detection and object tracking information are stored in a separate CSV file named “gt.csv”. The segmentation masks are stored in the folder “masks”. An additional folder is provided containing all images without ground truth annotations. The contents of the CSV file are explained in Table 6:

### Technical Validation

The growth experiments were conducted using multiple trays, resulting in different numbers of replicates per species. Here, a minimum of nine replicates were used. Seeds were treated according to breeders’ recommendations. Some weeds germinated only when treated with gibberellin acid. Therefore, the optimal procedure was evaluated in a prior experiment.

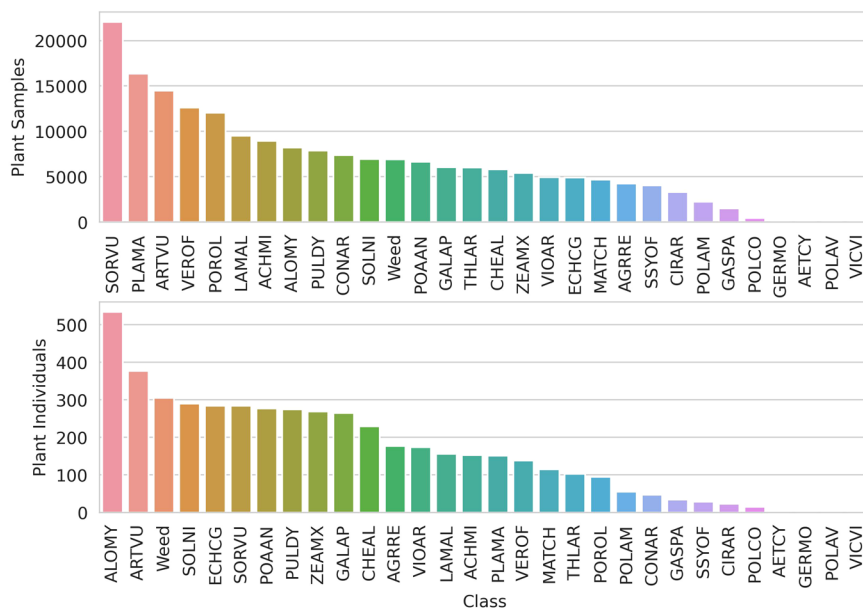
The quality of the bounding boxes and labels was ensured twofold. First, a valid bounding box could be evaluated during the labeling process by using tools in CVAT directly and by an additional human inspector doing quality control. Second, using the time-series information, all plant cut-outs of one plant individual could be plotted in a series of images to visually inspect the bounding boxes. Finally, plants of additional not sown species could be assessed and classified in an additional step. Remaining instances were labeled as class “Weed”, as they were tiny and thus could not be labeled by species.

The high variability in the seed germination rate of different weed species resulted in a very diverse data set of different weeds with different germination rates, as shown in Fig. 4.

### Usage Notes

The dataset can be downloaded via custom Python scripts (<https://github.com/grimmmlab/MFWD>) and used as a resource for precision agriculture, smart farming, and computer vision related tasks. In addition, we encourage the development of algorithms that take plant phenotyping data into account. Since the dataset consists of high-resolution images, we also provide a custom Python script to easily resize the images. The dataset could be a useful resource for the computer science community in general to develop novel machine learning and computer vision algorithms for automatic weed detection. Here, the data could be used in classification as well as object detection and segmentation tasks. Furthermore, the additional time series information makes our dataset suitable for multiple object tracking. Finally, the inherent inter- and intra-class variance can be used to evaluate new algorithms that address class imbalance, which is a challenge in many machine learning tasks<sup>31–33</sup>.

To demonstrate the usefulness of the dataset, we provide a simple baseline experiment on the image classification task. For this purpose, we use plant cut-outs from the jpeg-compressed images and rescale them to an image size of  $224 \times 224$  pixels<sup>2</sup>. Here we focused on the multi-species classification for sorghum, i.e. all sorghum



**Fig. 4** Distribution of plant samples (separate images) and plant individuals (multiple images of one plant through time) per class.

model	learning rate (*10 <sup>-4</sup> )	best epoch	f1-score (%)
EfficientNet_b0	2.37	8	88.75
EfficientNet_b0	8.93	1	83.93
<b>EfficientNet_b0</b>	<b>5.40</b>	<b>32</b>	<b>90.00</b>
EfficientNet_b0	3.97	8	89.35
EfficientNet_b0	1.43	26	88.88
ResNet-10	2.37	20	88.24
ResNet-10	8.93	1	81.28
ResNet-10	5.40	21	88.74
ResNet-10	3.97	20	88.70
ResNet-10	1.43	20	87.75

**Table 7.** Summary of the hyperparameter optimization calculated on the validation set.

varieties were labeled as “SORVU” and excluded all maize images. Additionally, the generic weed class was excluded for this experiment, as these were mostly small plants which could not even be classified by the human eye. We also excluded POLAV and VICVI from the experiment because they contained less than three plant individuals and thus could not be separated into training, validation, and test sets. We strongly recommend stratifying the data by plant individuals, as the temporal resolution is high, and models may overfit if the data were randomly split. The final dataset of 27 plant species contained 167,505 images and was split into a training- (~60%), validation- (~20%), and test-set (~20%).

We selected two different deep learning-based model architectures, ResNet-10<sup>34</sup> and EfficientNet\_b0<sup>35</sup>, to evaluate the classification performance. For the hyperparameter optimization we used grid-search and five different learning rates (sampled from a log uniform distribution in the range between 1e-3 and 1e-4). We sampled 512 plant cut-outs in a batch by oversampling the minority classes, due to high class imbalances. The networks were initialized with weights from ImageNet. The Adam<sup>36</sup> optimizer with a learning rate scheduler and the cross-entropy loss<sup>37</sup> was used to train the models. We validated the models using the validation-set by calculating the weighted f1-score<sup>38,39</sup>, due to the high class imbalance. Each model was trained for a maximum of 50 epochs. Early stopping<sup>40</sup> was used as a regularization technique to avoid overfitting. After training the models, EfficientNet\_b0 with a learning rate of  $\sim 5.4 \cdot 10^{-4}$  gave the best results on the validation set with an f1-score of 90.00%. The summary of the hyperparameter optimization is shown in Table 7.

Finally, the best performing model was applied to the hold-out test set to evaluate the generalization abilities on an unseen dataset. Here, the model achieved a weighted f1-score of 90.57%, indicating good generalization performance within the MFW dataset. The complete code for training and testing the model is publicly available in our GitHub repository.

However, deep learning models trained on this dataset may not be applicable to out-of-context data, such as weed detection in drone imagery. Here, pre-training a model on our dataset and fine-tuning it to the target



task might be a feasible strategy to scale up weed detection in agricultural landscapes. However, the main target application of our dataset is to encourage the research community to develop new computer vision algorithms on a unified dataset, thus increasing the reproducibility of the results.

### Code availability

The code to download the dataset is publicly available for download on GitHub: <https://github.com/grimmlab/MFWD>.

Received: 29 August 2023; Accepted: 10 January 2024;

Published online: 23 January 2024

### References

- Dentika, P., Ozier-Lafontaine, H. & Penet, L. Weeds as Pathogen Hosts and Disease Risk for Crops in the Wake of a Reduced Use of Herbicides: Evidence from Yam (*Dioscorea alata*) Fields and Colletotrichum Pathogens in the Tropics. *J Fungi* **7**, 283, <https://doi.org/10.3390/jof7040283> (2021).
- Norris, R. F. & Kogan, M. Interactions between weeds, arthropod pests, and their natural enemies in managed ecosystems. *Weed Sci* **48**, 94–158 (2000).
- Schumacher, M., Dieterich, M. & Gerhards, R. Effects of weed biodiversity on the ecosystem service of weed seed predation along a farming intensity gradient. *Glob Ecol Conserv*; **24**, e01316, <https://doi.org/10.1016/j.gecco.2020.e01316> (2020).
- Logsdon, S. D. Root effects on soil properties and processes: Synthesis and future research needs. In: *Enhancing Understanding and Quantification of Soil-Root Growth Interactions*. John Wiley & Sons, Ltd, pp 173–196, <https://doi.org/10.2134/advagricsystmodel4.c8> (2015).
- Harker, K. N., Clayton, G. W. & O'Donovan, J. T. Reducing agroecosystem vulnerability to weed invasion. In: *Invasive Plants: Ecological and Agricultural Aspects*. Birkhäuser Basel, pp 195–207, [https://doi.org/10.1007/3-7643-7380-6\\_12](https://doi.org/10.1007/3-7643-7380-6_12) (2005).
- Harker, K. N. & O'Donovan, J. T. Recent Weed Control, Weed Management, and Integrated Weed Management. *Weed Technol*; **27**, 1–11, <https://doi.org/10.1614/WT-D-12-00109.1> (2013).
- Myers, J. P. *et al.* Concerns over use of glyphosate-based herbicides and risks associated with exposures: A consensus statement. *Environ. Heal. A Glob. Access Sci. Source*. **15**, 1–13, <https://doi.org/10.1186/s12940-016-0117-0> (2016).
- Steinmetz, Z. *et al.* Plastic mulching in agriculture. Trading short-term agronomic benefits for long-term soil degradation? *Sci. Total Environ.*; **550**, 690–705, <https://doi.org/10.1016/j.scitotenv.2016.01.153> (2016).
- MacLaren, C., Storkey, J., Menegat, A., Metcalfe, H. & Dehnen-Schmutz, K. An ecological future for weed science to sustain crop production and the environment. A review. *Agron. Sustain. Dev.* **40**, 1–29, <https://doi.org/10.1007/s13593-020-00631-6> (2020).
- Christensen, S. *et al.* Site-specific weed control technologies. *Weed Res*; **49**, 233–241, <https://doi.org/10.1111/j.1365-3180.2009.00696.x> (2009).
- Hasan, A. S. M. M., Sohel, F., Diepeveen, D., Laga, H. & Jones, M. G. K. A survey of deep learning techniques for weed detection from images. *Comput. Electron. Agric.*; **184**, <https://doi.org/10.1016/j.compag.2021.106067> (2021).
- Wang, A., Zhang, W. & Wei, X. A review on weed detection using ground-based machine vision and image processing techniques. *Comput. Electron. Agric.*; **158**, 226–240, <https://doi.org/10.1016/j.compag.2019.02.005> (2019).
- Dian Bah, M., Hafiane, A. & Canals, R. Deep learning with unsupervised data labeling for weed detection in line crops in UAV images. *Remote Sens*; **10**, <https://doi.org/10.3390/rs10111690> (2018).
- Bakhshpour, A. & Jafari, A. Evaluation of support vector machine and artificial neural networks in weed detection using shape features. *Comput. Electron. Agric.*; **145**, 153–160, <https://doi.org/10.1016/j.compag.2017.12.032> (2018).
- Sivakumar, A. N. V. *et al.* Comparison of object detection and patch-based classification deep learning models on mid-to late-season weed detection in UAV imagery. *Remote Sens*; **12**, <https://doi.org/10.3390/rs12132136> (2020).
- Genze, N. *et al.* Improved weed segmentation in UAV imagery of sorghum fields with a combined deblurring segmentation model. *Plant Methods*; **19**, <https://doi.org/10.1186/s13007-023-01060-8> (2023).
- Genze, N. *et al.* Deep learning-based early weed segmentation using motion blurred UAV images of sorghum fields. *Comput. Electron. Agric.*; **202**, <https://doi.org/10.1016/j.compag.2022.107388> (2022).
- Milioto, A., Lottes, P. & Stachniss, C. Real-Time Semantic Segmentation of Crop and Weed for Precision Agriculture Robots Leveraging Background Knowledge in CNNs. In: *Proceedings - IEEE International Conference on Robotics and Automation*. Institute of Electrical and Electronics Engineers Inc., pp 2229–2235, <https://doi.org/10.1109/ICRA.2018.8460962> (2018).
- Lottes, P., Khanna, R., Pfeifer, J., Siegart, R. & Stachniss, C. UAV-based crop and weed classification for smart farming. In: *Proceedings - IEEE International Conference on Robotics and Automation*. Institute of Electrical and Electronics Engineers Inc., pp 3024–3031, <https://doi.org/10.1109/ICRA.2017.7989347> (2017).
- Zhao, Z. Q., Zheng, P., Xu, S. T. & Wu, X. Object Detection with Deep Learning: A Review. *IEEE Trans. Neural Networks Learn. Syst.*; **30**, 3212–3232, <https://doi.org/10.1109/TNNLS.2018.2876865> (2019).
- Genze, N., Bharti, R., Grieb, M., Schultheiss, S. J. & Grimm, D. G. Accurate machine learning-based germination detection, prediction and quality assessment of three grain crops. *Plant Methods*; **16**, <https://doi.org/10.1186/s13007-020-00699-x> (2020).
- Janneh, L. L., Zhang, Y., Cui, Z. & Yang, Y. Multi-level feature re-weighted fusion for the semantic segmentation of crops and weeds. *J King Saud Univ - Comput Inf Sci*; **35**, <https://doi.org/10.1016/j.jksuci.2023.03.023> (2023).
- Long, J., Shelhamer, E. & Darrell, T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, pp 431–440, <https://doi.org/10.1109/CVPR.2015.7298965> (2015).
- Champ, J. *et al.* Instance segmentation for the fine detection of crop and weed plants by precision agricultural robots. *Appl Plant Sci*; **8**, <https://doi.org/10.1002/aps3.11373> (2020).
- von Redwitz, C. *et al.* Better-Weeds – Next generation weed management. *Tagungsband 30 Dtsch Arbeitsbesprechung über Frag der Unkrautbiologie und -bekämpfung*; 432–437, <https://doi.org/10.5073/20220124-075254> (2022).
- Sapkota, B. B. *et al.* Use of synthetic images for training a deep learning model for weed detection and biomass estimation in cotton. *Sci Rep*; **12**, <https://doi.org/10.1038/s41598-022-23399-z> (2022).
- Kirillov, A., He, K., Girshick, R., Rother, C. & Dollar, P. Panoptic segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE Computer Society, pp 9396–9405, <https://doi.org/10.1109/CVPR.2019.00963> (2019).
- Madsen, S. L. *et al.* Open plant phenotype database of common weeds in Denmark. *Remote Sens*; **12**, <https://doi.org/10.3390/RS12081246> (2020).
- Garcin, C., Joly, A., Bonnet, P., Chouet, M. & Servajean, M. Pl@ntNet-300K: a plant image dataset with high label ambiguity and a long-tailed distribution. *NeurIPS 2021-35th Conf Neural Inf Process Syst* 2021.
- Genze, N. *et al.* Manually annotated and curated dataset of diverse weed species in maize and sorghum for computer vision. *Technical University of Munich, mediaTUM*, <https://doi.org/10.14459/2023mp1717366> (2023).

31. Chou, H. P., Chang, S. C., Pan, J. Y., Wei, W. & Juan, D. C. Remix: Rebalanced Mixup. In: *Lecture Notes in Computer Science*. Springer Science and Business Media Deutschland GmbH, pp 95–110, [https://doi.org/10.1007/978-3-030-65414-6\\_9](https://doi.org/10.1007/978-3-030-65414-6_9) (2020).
32. Zhou, F., Yang, S., Fujita, H., Chen, D. & Wen, C. Deep learning fault diagnosis method based on global optimization GAN for unbalanced data. *Knowledge-Based Syst*; **187**. <https://doi.org/10.1016/j.knsys.2019.07.008> (2020).
33. Kaur, H., Pannu, H. S. & Malhi, A. K. A systematic review on imbalanced data challenges in machine learning: Applications and solutions. *ACM Comput. Surv.*; **52**. <https://doi.org/10.1145/3343440> (2019).
34. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp 770–778, <https://doi.org/10.1109/CVPR.2016.90> (2016).
35. Tan, M. & Le, Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks. In: *36th International Conference on Machine Learning, ICML 2019*. 2019, pp 10691–10700.
36. Kingma, D. & Ba, J. Adam: A Method for Stochastic Optimization. In: *International Conference on Learning Representations (ICLR)*. San Diego, CA, USA, 2015.
37. Cox, D. R. The Regression Analysis of Binary Sequences. *J R Stat Soc Ser B* **20**, 215–232, <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x> (1958).
38. Chinchor, N. MUC-4 evaluation metrics. *4th Messag Underst Conf MUC 1992 - Proc*: 22–29, <https://doi.org/10.3115/1072064.1072067> (1992).
39. He, H. & Garcia, E. A. Learning from Imbalanced Data. *IEEE Trans Knowl Data Eng* **21**, 1263–1284, <https://doi.org/10.1109/TKDE.2008.239> (2009).
40. Yao, Y., Rosasco, L. & Caponnetto, A. On early stopping in gradient descent learning. *Constr Approx* **26**, 289–315, <https://doi.org/10.1007/s00365-006-0663-2> (2007).
41. Hughes, D. P. & Salathe, M. An open access repository of images on plant health to enable the development of mobile disease diagnostics. *arXiv.org* 2015. <http://arxiv.org/abs/1511.08060> (accessed 27 Oct2023).
42. Jiang, Y., Li, C., Paterson, A. H. & Robertson, J. S. DeepSeedling: Deep convolutional network and Kalman filter for plant seedling detection and counting in the field. *Plant Methods*; **15**. <https://doi.org/10.1186/s13007-019-0528-3> (2019).
43. Olsen, A. *et al.* DeepWeeds: A Multiclass Weed Species Image Dataset for Deep Learning. *Sci Rep*; **9**. <https://doi.org/10.1038/s41598-018-38343-3> (2019).
44. Wang, P. *et al.* Weed25: A deep learning dataset for weed identification. *Front Plant Sci*; **13**. <https://doi.org/10.3389/fpls.2022.1053329> (2022).
45. Kitzler, F., Barta, N., Neugschwandtner, R. W., Gronauer, A. & Motsch, V. WE3DS: An RGB-D Image Dataset for Semantic Segmentation in Agriculture. *Sensors*; **23**. <https://doi.org/10.3390/s23052713> (2023).
46. Sudars, K., Jasko, J., Namatevs, I., Ozola, L. & Badaukis, N. Dataset of annotated food crops and weed images for robotic computer vision control. *Data Br*; **31**. <https://doi.org/10.1016/j.dib.2020.105833> (2020).
47. Gisellsson, T. M., Jørgensen, R. N., Jensen, P. K., Dyrmann, M. & Midtiby, H. S. A Public Image Database for Benchmark of Plant Seedling Classification Algorithms. *arXiv.org* 2017. <http://arxiv.org/abs/1711.05458> (accessed 27 Oct2023).
48. Bender, A., Whelan, B. & Sukkarieh, S. A high-resolution, multimodal data set for agricultural robotics: A Ladybird’s-eye view of Brassica. *J F Robot*; **37**, 73–96, <https://doi.org/10.1002/rob.21877> (2020).
49. Haug, S. & Ostermann, J. A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In: *Lecture Notes in Computer Science*. Springer Verlag, pp 105–116, [https://doi.org/10.1007/978-3-319-16220-1\\_8](https://doi.org/10.1007/978-3-319-16220-1_8) (2015).
50. Minervini, M., Fischbach, A., Scharr, H. & Tsafaris, S. A. Finely-grained annotated datasets for image-based plant phenotyping. *Pattern Recognit Lett*; **81**, 80–89, <https://doi.org/10.1016/j.patrec.2015.10.013> (2016).

## Acknowledgements

We thank Marlene Hanly and Agnes Wappler for their assistance in labeling the dataset. Funding for the research presented in this paper is provided by the Bavarian State Ministry for Food, Agriculture and Forests within the EWIS project (Funding ID: G2/N/19/13).

## Author contributions

D.G.G., N.G., M.G. conceived and designed the study. J.G. and W.K.V. acquired the plants’ seeds. J.G. and W.K.V. designed the experiments and were responsible for the image acquisition. N.G. and M.W. labeled the dataset. M.G. guided the labeling process with domain expertise. N.G. was responsible for quality control of the data. N.G. and D.G.G. wrote the manuscript with contributions from all authors. All authors read and approved the final manuscript.

## Funding

Open Access funding enabled and organized by Projekt DEAL.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to D.G.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024