Deep-learning-ready RGB-depth images of seedling development

Félix Mercier, Geoffroy Couasnet, Angelina El Ghaziri, Nizar Bouhlel, Alain Sarniguet, Muriel Marchi, Matthieu Barret, David Rousseau

Plant Methods

DATABASE

Open Access

Deep-learning-ready RGB-depth images of seedling development

Félix Mercier[1], Geoffroy Couasnet[1], Angelina El Ghaziri[2,3], Nizar Bouhlel[2,3], Alain Sarniguet[1,3,4], Muriel Marchi[1,3,4], Matthieu Barret[1,3,4] and David Rousseau[1,4*]

Abstract

In the era of machine learning-driven plant imaging, the production of annotated datasets is a very important contribution. In this data paper, a unique annotated dataset of seedling emergence kinetics is proposed. It is composed of almost 70,000 RGB-depth frames and more than 700,000 plant annotations. The dataset is shown valuable for training deep learning models and performing high-throughput phenotyping by imaging. The ability of such models to generalize to several species and outperform the state-of-the-art owing to the delivered dataset is demonstrated. We also discuss how this dataset raises new questions in plant phenotyping.

Keywords RGB-depth, Seedling kinetics, Deep learning, Data set

Background

Plants continuously metamorphose throughout their development, with specific stages occurring in a predetermined order. This paper introduces a dataset on key seedling development stages from germination to first leaf formation. During this period, plants must adapt to their environment for successful photosynthesis and growth. Seedlings undergo photomorphogenesis upon emerging from the soil, which involves reduced hypocotyl growth, cotyledon opening, photosynthesis initiation, and meristem activation. These processes require extensive genomic reprogramming and are challenging to study owing to seedling population variability. Understanding seedling development is essential for determining crop yield, as uneven emergence can lead to lower yields and poor farmer acceptance. More precisely, the kinetics of seedling development is a measure of seed vigor which is an important agricultural trait defined by (i) rapid and uniform germination and (ii) seedling growth, that is determined by physiological and sanitary quality [1]. While germination kinetics can be monitored in soilless systems [2], measuring seedling growth under real growth conditions (i.e. with a culture substrate) is not yet possible. The method proposed in this data paper enables the measurement of seedling growth in a high throughput manner with fine temporal resolution. This approach could be used to assess the vigor of seed lots from different genotypes or varieties, under both suboptimal and optimal conditions (e.g. temperature, humidity) and during exposure with a diversity of microorganisms.

High-throughput kinetic monitoring of seedling growth has been performed on RGB-depth images in recent studies of Samiei et al. [3], Garbouge et al. [4], and Couasnet et al. [5]. Color imaging was first successfully coupled with deep learning models by Samiei et al. [3] to monitor*Medicago*. The added value of fusing RGB with depth was introduced by Garbouge et al. [4] to monitor several varieties of beans. To reproduce such experiments, a software was developed and provided to the scientific community by Couasnet et al. [5]. In this data

*Correspondence:
David Rousseau
david.rousseau@univ-angers.fr
[1] Université d'Angers, 40 Rue de Rennes, 49000 Angers, France
[2] Institut Agro, 2 rue André Le Nôtre, 49000 Angers, France
[3] UMR1345, Institut de Recherche en Horticulture et Semences (IRHS), 49071 Beaucouzé, France
[4] INRAE, 42 Rue Georges Morel, 49071 Beaucouzé, France

paper, we provide a dataset of RGB depth monitoring data of seedling kinetics extended to more plant species. This dataset is annotated and functional for deep learning use, i.e., it is reusable for scientists who are eager to train new deep learning architectures.

We demonstrate the value of the delivered dataset by training a new deep learning model that shows good generalization performance on various species in comparison with the types of models optimized for single species by Samiei et al. [3], Garbouge et al. [4], and Couasnet et al. [5]. Finally, we discuss the perspectives of the available research owing to the accessibility of this data set.

Construction and content

Room and equipment

Data acquisition was performed via eight Intel RealSense D435 cameras, strategically positioned above two tables on either side of the room. These cameras were interconnected with a Raspberry Pi 3 model B, referred to as the *clients*, where RGB, depth, and infrared images were temporarily stored. At hourly intervals, each of the eight *clients* transmitted the images via a local Wi-Fi network to a ninth Raspberry Pi 3 model B, identical to the others, designated as the *server*, which was equipped with a 4TB external hard drive. After image transmission, the sizes of the original images were compared with those of the transmitted images, and if they matched, the original images were deleted from the clients. A system monitoring script on the server regularly verified the integrity of data acquisition, and in the event of any issues, a warning email was sent to maintenance. Each *client* was also equipped with Power over Ethernet (*PoE*), enabling the Raspberry Pis to be powered via Ethernet cables, and a 64 GB Samsung microSDXC card, allowing for continued data acquisition during network disconnection (with a capacity of several days). Potential power interruptions were mitigated by the presence of an uninterruptible power supply (UPS) capable of providing backup power for 30 min to 1 h. Last, the *clients* and *server* were interconnected on the same local area network, with Wi-Fi connectivity for the *server* and wired connections for the *clients*, linked to a TP-Link T1500-28PCT switch and administered by a TP-Link Archer C9 router. The hardware components necessary for installing an 8-camera acquisition system in a controlled environment are listed in Table 1.

The controlled environment growth chamber was equipped with eight Intel RealSense D435 cameras, periodically positioned across two rows on either side of the room. The cameras were not always operational simultaneously, and the number and format of germination trays positioned beneath each camera varied. Two types of trays were used: 40-pot trays (5 columns × 8 rows) and 84-pot trays (7 columns × 12 rows). A detailed description of each experiment is provided in the subsequent section. To provide explicit insight into the data acquisition environment, a map of the experiment, including the corresponding camera designations, is presented in Fig. 1.

Table 1 List of equipment used to equip the growth chamber

Category	Name and model	Number	Details
UPS	APC Back-UPS CS	1	600VA, 400W
Router	TP-Link Archer C9	1	v5.0
Switch	TP-Link T1500-28PCT	1	24 × 10/100Mbps × RJ45 Ports
Fixing bar	Aluminium section Bosch rexroth	2	30 × 30 × 3000 mm
Sensors	Intel Realsense D435	8	RGB, Depth, IR
Nano-Computer	Raspberry Pi 3 Model B/4	9	4Go RAM
Raspberry Case	Official case For Raspberry Pi 3B		
Power Supply	Raspberry Pi PoE+ HAT		Power supply via Ethernet
Connections	Category 5e S/FTP Roline ethernet cable (Black)	8	10m with Connections
Memory card	Samsung microSDXC 64Go	9	
External hard drive	LaCie USB-C 4To	1	Hard disk 2.5″

Although images are acquired every 15-minutes, the acquisition scripts are designed to activate the camera sensors 5 min before acquisition. The images are subsequently taken and saved, and the sensors are then deactivated for approximately 10 min before the next acquisition cycle. A detailed explanation of this process can be found in Garbouge et al. [4]. In the meanwhile, a data transfer script exports the images to an external hard drive connected to the *server*. To compensate for data loss resulting from network interruptions, power failures, or hardware malfunctions, we have developed a script to detect troubleshooting in the arrival of data on the *server*. This script verifies that images are received by the *server* at hourly intervals (with an additional check performed an hour later if necessary) and that the image size exceeds a predetermined threshold (corresponding to an empty image of identical size and format). If any of these conditions are not met, an alert is triggered, and an email is sent detailing the problem(s) identified and the hardware concerned. Additionally, a Raspberry Pi installed in the offices performs hourly checks to ensure that the room's server is accessible. This supplementary security

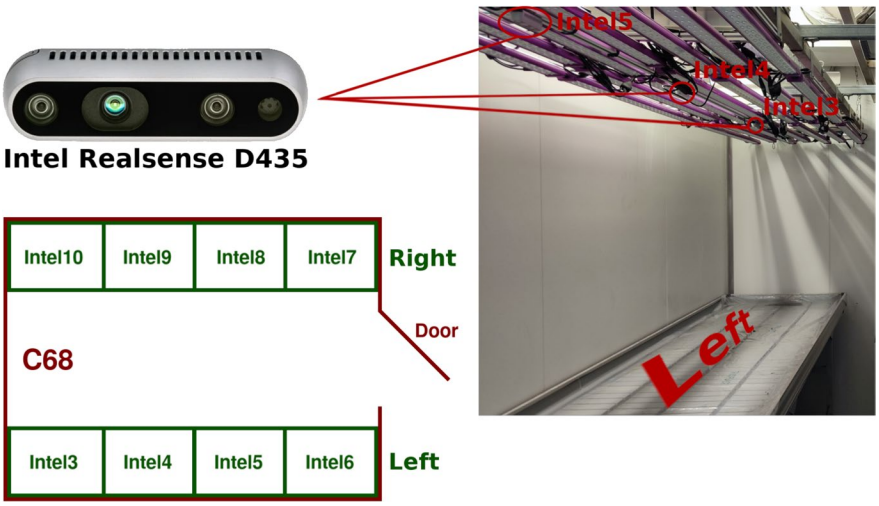

Fig. 1 Detailed visualisation of the experimental setup used to produce the delivered dataset. A RGB-Depth *Intel Realsense D435* camera (top left). Cameras positioned above the plants for top view monitoring (top right). The cameras are labeled with a numerical identifier ranging from 3 to 10. The chamber map is shown at the bottom left

measure prevents emails alert system failures in the event of an internet access interruption. Despite the implementation of these security systems (UPS, acquisition verification script, and server accessibility verification), we were unable to cope with power interruptions exceeding 1 h or system shutdowns occurring during nighttime or public holidays. Consequently, image acquisition may have been interrupted on several occasions, resulting in potential gaps in the datasets provided. Nevertheless, it is important to note that the acquisition time remains accurate. Therefore, spurious missing data do not affect the timeline. The data loss is estimated to be around % which is considered negligible. We assume that the missing data are uniformly distributed over the time and do not interfere with the biological process being monitored.

Intel RealSense D435 cameras include RGB camera, two infrared (IR) cameras, and an IR projector. The product data sheet [6] provides a detailed explanation of how the depth image is calculated from the stereo IR sensor. The IR projector enhances image quality on surfaces with minimal texture. The pixel values in the depth image represent distances to the camera in millimeters. Notably, the depth field exceeds the color field in terms of spatial extent. The raw resolution of the images is 1920×1080 pixels for RGB and 1280×720 pixels for depth. To facilitate alignment of the RGB and depth images, the depth image undergoes cropping and resizing via nearest neighbor interpolation. This transformation is performed via the *PyRealSense2* package associated with the camera, which relies on intrinsic camera parameters. To provide readily usable data, the dataset includes the *RGB full frame* and the *aligned depth frame*. The concatenation of these four channels is referred to as an *RGB-depth full frame*. As illustrated in Fig. 2, the *RGB full frame* presents a top view of the

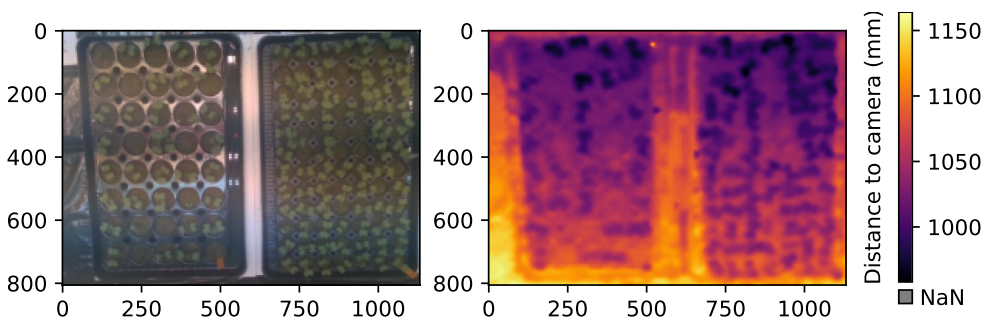

Fig. 2 Cropped *RGB full frame* (left) and *Depth aligned frame* (right). The images are cropped to exclude distant objects and emphasize depth dynamics

tray, enabling visualization of all the plants with their distance-to-camera values linked to their height.

Data description

The available database comprises multiple sequences of *RGB-depth full frames*, collectively referred to as *full time lapses*. In contrast, we define *pots time lapses* as the cropped pot regions extracted from full time lapses. Throughout the time lapse sequences, we observed the growth and development of the plants, which can be described by biologists in distinct developmental stages. Four main stages of development can be observed in the *pots time lapses*. As illustrated in Fig. 3, this inclues the soil (when the seedling is still fully in the soil), first appearance of the cotyledons, opening of the cotyledons and appearance of the first leaf. The dataset includes annotations for a set of dicotyledon species, which were manually obtained through collaborative efforts with biologists.

The datasets presented in this article include *full time lapses* of various plant species, including rapeseed, tomatoes, and beans, acquired during the monitoring of 11 distinct trials conducted throughout 2022. A comprehensive list of these 11 trials, accompanied by relevant information (camera specifications, tray configurations, plant varieties, germination pot details, start and end dates, and image acquisition quantities), is provided in the Appendix section. Phenotyping was performed on beans, rapeseed, and tomato species. A summary of the phenotyped plants per species is presented in Table 2.

The dataset is publicly accessible in the DATA INRAE repository, DOI: https://doi.org/10.57745/AMFJTK. The file tree structure is illustrated in Fig. 4. The dataset is organized into 11 compressed *.zip* files, each corresponding to a distinct trial. Within these files, images are sorted chronologically by acquisition start date, then by camera, and stored in *.png* format within dedicated color and depth folders. Labels are also stored in *.zip* files, sorted by trial and camera, with one file per plant.

Table 2 Synthesis of the *full time-lapse* and *RGB-Depth full frame* quantity per species

Species	Pots time-lapse	Labelled pots time-lapse	RGB-depth full frame
Rapeseed	1 760	336	15 218
Tomatoes	1 960	480	33 283
Beans	2 320	400	21 445
Total	6 040	1 216	69 946

The corresponding *.xlsx* tables provide a summary of the image and label contents. General information about the dataset is provided in the *Readme.txt* file.

We now position the provided dataset with the related literature. There are currently many publicly available plant datasets, as reviewed by Kurtser and Lowry [7], Das Choudhury et al. [8], Lu and Young [9] and Cao et al. [10]. In this study, we provide a multispecies, indoor, multimodal RGB-depth time-lapse (15-minute frame rate) dataset annotated for the classification of seedling growth stages from a top view. The main features of the most related datasets and our dataset are provided in Table 3. In Cruz et al. [11], authors operated with two species (including beans as in our dataset) viewed from the top view in time lapse (1-hour frame rate) of RGB-depth images. This dataset also includes fluorescence and infrared but provides data for only 21 plants. Additionally, [12] operated with two species from a top view in time lapses acquired at a frame rate of 20 min. This dataset includes only RGB data and is limited to 147 plants. Similarly, [13] provides RGB top view images. Only one species is studied and 200 plants are phenotyped with a very low frame rate (18 days). The dataset is designed to perform vegetative and reproductive stage classification.

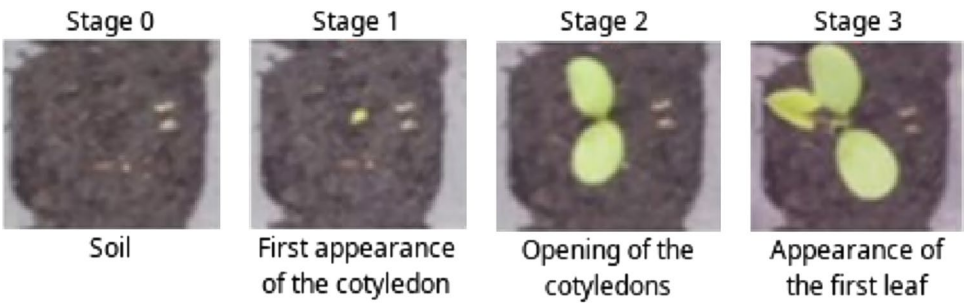

Fig. 3 RGB images extracted from a *pots time-lapse* at each stage

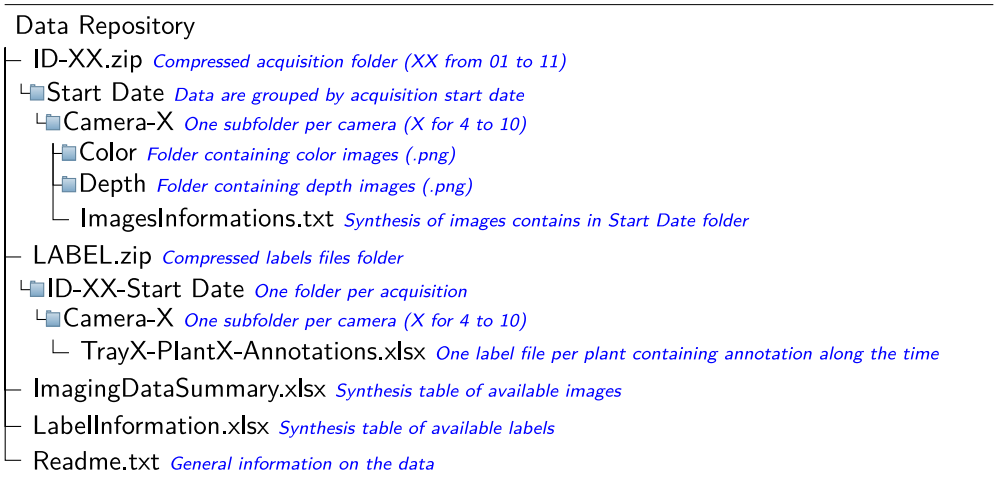

Fig. 4 Data repository tree and comments available in the DATA INRAE repository, DOI: https://doi.org/10.57745/AMFJTK

Table 3 Comparison of key characteristics among datasets from previous most related studies and our dataset

Data paper References	Species	Plants	Image Modality	Acquisition Frequency
Cruz et al. [11]	*Arabidopsis*, Bean	21	RGB-Depth, IR, fluo	1 hour
Scharr et al. [12]	*Arabidopsis*, Tobacco	147	RGB	From 20 minutes to 6 hours
Gené-Mola et al. [15]	Apple tree	12 839*	RGB-Depth, IR	No time-lapse
Uchiyama et al. [14]	Komatsuna	5	RGB-Depth	4 hours
Lac et al. [16]	Maize, Bean, Leek	8247	RGB	No time-lapse
Genze et al. [17]	28 weed species	5000	RGB	2 images per day
Liang et al. [18]	Maize	156	RGB, IR, fluo, HS	Daily
Arend et al. [19]	*Arabidopsis*	484	RGB, IR, fluo	Daily
Arrechea-Castillo et al. [20]	*Urochloa*	200	RGB	≈ 18 days
Our dataset	Rapeseed,Tomato, Bean	6040	RGB-depth	15 minutes

HS stands for HyperSpectral, IR for infra-red and fluo for chlorophyl fluorescence imaging. The number marked with * is fruit number instead of plants number

Finally [14], provided RGB-depth from the top view, with acquisition every 4 h. This dataset is dedicated to leaf segmentation; however, it contains only 5 plants and is limited to only one type of plant, Komatsuna.[1]

The number of phenotyped plants is crucial for capturing the biological diversity. We achieved a dataset size comparable in magnitude to the largest dataset reported in the literature. This enhances the robustness of the knowledge to be extracted from this dataset. The largest data sets reported by Gené-Mola et al. [15], Lac et al. [16], and Genze et al. [17] do not involve time-lapse imaging, and/or rely solely on RGB data. This strongly differs from our multimodal, time-resolved approach which appears as an original dataset in comparison with the most related publicly available datasets.

Utility and discussion

We now detail the utility of the delivered dataset and discuss further developments opened with this dataset.

Pot scale classification

Before processing, the *Pots frames* undergo normalization. The RGB channels, which are represented as 8-bit integer data, are divided by 255. The depth channel contains missing values, which are filled by inpainting via anisotropic diffusion. A Laplacian filter is applied iteratively, and through diffusion, missing values are progressively replaced. The depth pixels, represented as 16-bit integers, are divided by 65,535. This normalization between 0 and 1 preserves precision. The RGB and depth

[1] Japanese mustard spinach.

ranges are the same to facilitate weight optimization during deep learning training.

The software *GrowthData* [5] was provided with a model trained on images of Bean. With the dataset provided in this article, we trained a multi-species model. The annotated data set was split into training, validation, and testing sets, with the following ratio 80%, 10%, and 10%, respectively. Because of the similarity between successive *pot frames*, a *pot time lapse* lies entirely within one of the three subsets. Detailed information regarding the number of Pots time lapses and frames is provided in Table 4.

The multispecies model (MS model) was trained on the entire training dataset. To increase the variability in the training set, classical data augmentation techniques, including rotation (90°, 180°, or 270°), and vertical and horizontal symmetry, were employed. For comparison, we trained the Rp model via rapeseed data, the T model via tomato data, and the B model via Bean data. All the models utilized the architecture of the original convolutional neural network (CNN) of Garbouge et al. [4] presented in Fig. 5. The model includes four convolution blocks with 32, 64, 128, and 128 filters; a fully connected layer with 256 neurons; a dropout with a drop rate of 0.5; and a final fully connected layer with four outputs, corresponding to the number of classes. Each convolution block performs 3×3 pixel filtering, activation, and 2×2 pixel max pooling. Following convolution and a fully connected layer, all the activations were rectified linear units (ReLUs), except for the last one, which was a Softmax function.

The CNN we designed serves as a simple baseline approach. It predicts the class of each frame in the time-lapse independently, without considering the order in which the frames occur. Given a training set of N pairs of images x_i and labels y_i, we trained the parameters θ of the network f using stochastic gradient descent to minimize empirical risk:

$$\theta^* = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^{N} \mathcal{L}(y_i, f(x_i, \theta)), \tag{1}$$

where $\mathcal{L}$ denotes the loss function chosen as the categorical cross entropy. This loss function is defined in Demirkaya et al. [21] as

$$\mathcal{L}(y_i, f(x_i, \theta)) = -\sum_{k=1}^{4} \delta(y_i - k) \log(f_k(x_i, \theta)), \tag{2}$$

where k is the class index, $\delta(y_i - k)$ is equal to 1 if $y_i = k$ and 0 otherwise and $f_k(x_i, \theta)$ denotes the probability that x_i belongs to class k. The gradient descent aims the minimization using the Adam optimizer algorithm introduced in Kingma and Ba [22]. We adopted an adaptive learning rate. If the validation loss does not improve for 10 consecutive epochs, the learning rate is halved. For

Table 4 Details of the data split per species in terms of *Pots time-lapses* and *Pots frame*

Species	Pots time-lapses			Pots frames		
	Training	Validation	Testing	Training	Validation	Testing
Rapeseed	268	34	34	214 880	27 795	27 431
Tomato	384	48	48	244 345	30 467	31 108
Bean	320	40	40	148 864	18 608	18 608
Total	972	122	122	608 089	76 870	77 147

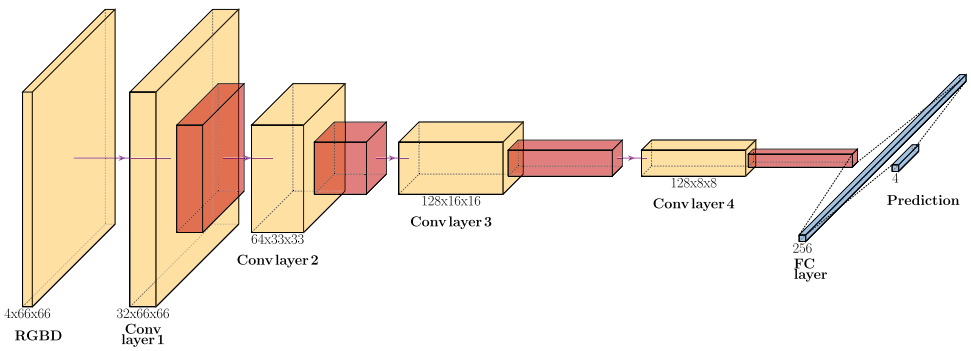

Fig. 5 Convolutional neural network used as a baseline to perform growth stage classification from RGB depth images. Orange, red, and blue blocks represent convolution, maximum pooling, and fully connected layers, respectively

training the T, B, and Rp models, the initial learning rate is 6.10^{-5}, while for training the MS model, it is 3.10^{-5}. The batch size is 2048 for training the MS, T, and B models, and 4096 for the Rp model.

The class frequency in the training set depends on the stage duration. Images within a time series are classified independently, which may not accurately predict the evolution of growth stages. To address this limitation, predictions are post-processed with a smoothing filter based on a sliding window computing a majority voting by finding the median of classes

$$s = \lfloor * \rfloor \left\{ \left(\frac{n+1}{2} \right)^{th} \right\}, \tag{3}$$

where s represent the smoothed predictions and n is the window's length (taken as 5 in this study). Then, smoothed class is forced to be uniform between the first and last appearance of a stage.

We evaluated the raw predictions via loss and accuracy (Rw-Acc). Additionally, we evaluated the corrected predictions in terms of accuracy over rapeseed data (Rp-Acc), tomato data (T-Acc), bean data (B-Acc), and overall data (O-Acc). The training process was repeated 10 times, and we computed the average performance. The results are presented in Table 5. The indicated uncertainty represents the 95% confidence interval computed from the standard deviation under the Gaussian hypothesis.

With respect to accuracy per species, each test set is best predicted by its respective specific model. The Rp model achieves an accuracy of 89.2% for rapeseed, the T model 87.9% for tomatoes, and the B model 94.8% for beans. The application of mono-species models to other species yields unsatisfactory results. The most notable example of transferability is the B model on tomato data, with an accuracy of 68.5%. In other cases, the accuracy does not exceed 57.2%. Across all species, the MS model reached an overall accuracy of 88.7%. The MS model performs slightly less accurately for tomatoes, with an accuracy of 88.3%, and more accurately for beans, with an accuracy of 89.7%. Although the global performance of the MS model is lower than that of each mono-species model, this represents a gain in generality. By acquiring a large, mixed-species dataset, we developed a model capable of processing multiple species with a comparable level of accuracy. This demonstrates the added value of the dataset delivered.

Exploring research trends

Several other uses of the delivered dataset can be envisioned, as detailed in this section. First, one can seek to outperform the baseline we provided in the previous section for the classification of seedling development stages. The way we perform individual *pot frame* classification with a memoryless convolutional neural network can be seen as a brute-force approach since it does not consider any prior information on plant development. There is, of course, a sequential causal aspect in plant growth, where each stage is intrinsically dependent on the preceding stages. An alternative neural network architecture that would incorporate this prior knowledge would be expected to provide higher performance than the architecture delivered with this data paper. One could consider recurrent neural networks (RNNs) or long short-term memory (LSTM) networks to model temporal dependencies. However, as demonstrated by Garbouge et al. [23], this introduces a dependency on the speed of plant growth and the frame rate of data capture. Faster-growing plants or higher frame rate data may skew the learning algorithms, introducing inaccuracies in the phenotypic analysis. Another approach would be to use ordinal loss functions, as proposed by Cao et al. [24], and Shi et al. [25]. The ordinal loss functions assign different penalties to errors based on their ordinal relationships, thus maintaining some degree of temporal causality. Such new investigations are now accessible to the reader via the dataset delivered in this article.

The dataset was produced with periodic sampling. However, one could be interested in acquiring data at non-periodic sampling rate. Different existing approaches [26, 27] for non-periodic sampling strategies could be tested indeed to reduce the cost of data storage [28, 29]. While the real interest would be to perform such non periodic sampling on-the-fly, it is possible to

Table 5 Average performance of the model trained with different species in terms of loss and accuracy

	Loss	Rw-Acc	O-Acc	Rp-Acc	T-Acc	B-Acc
Rp model	1.20 ± 0.89	70.1 ± 12.1	65.3 ± 8.1	89.2 ± 0.9	52.9 ± 14.4	46.8 ± 10.6
T model	0.96 ± 0.08	70.1 ± 2.0	69.1 ± 5.1	57.2 ± 12.2	87.9 ± 2.0	51.5 ± 3.5
B model	1.81 ± 0.27	61.8 ± 1.8	67.7 ± 2.6	52.6 ± 5.2	68.5 ± 2.0	94.8 ± 0.5
MS model	0.32 ± 0.01	88.3 ± 0.5	88.7 ± 1.3	88.7 ± 1.0	88.3 ± 1.8	89.7 ± 3.2

The associated uncertainty with a 95%-confidence interval over ten repetitions. Accuracy is given in percentage

simulate and investigate non-periodic sampling strategies via sub-sampling with the dataset provided thanks to the relatively high frame rate used.

Another limitation of classifying plant growth at the pot scale is the issue of overlapping plants. As the duration of data acquisition increased, the size of the seedlings exceeded the size of the pots, and occlusions of the plants appeared. Also, such occlusions would be likely to appear in the study of mixture of species which is important in agroecology. Advanced techniques for instantaneous leaf segmentation have shown promise in initiating occlusion detection [30, 31]. Occlusion detection enables the system to either stop processing or adopt a different processing method, ensuring classification performance. To extend the processing time, some methods exploit the circadian cycle to capture unobstructed frames [32] without any tracking method. To allow further investigation, including other tasks than classification (including tracking, segmentation, object detection, ...), we provide the *full frames* in the dataset.

Moreover, since we have described how to reproduce the data acquisition process in this article, we allow the reader to reproduce similar experiments with other species. One open question would then be how to adapt the already trained model with our dataset to new species. The dataset comprises three species. This already allows researchers to simulate such situations. For example, one can explore model adaptation techniques (domain adaptation, fine-tuning, transfer learning).

Finally, the provided dataset can serve any deep learning methodological task. To point a few, this could include neural architecture search [33], distillation knowledge to propose deep learning models that could run on the edge [34], i.e. on the minicomputer of the imaging system. One could also think of reducing the complexity of the dataset via data distillation [35] or the use of hybrid agronomical models that mixes mechanistic and machine learning approaches [36, 37]. As a last interesting perspective opened with the provided dataset, one could think of using the RGB-Depth images to serve other common plant imaging modalities (thermography, chlorophyll fluorescence, spectral imaging, ...) via style transfer [38, 39].

Conclusion

The dataset proposed in this data paper provides, to the best of our knowledge, the largest annotated dataset (70,000 frames and 700,000 annotations) on multi-species seedling development when viewed indoors via RGB-depth in time-lapse from a top view. The dataset has shown interest by comparison with the most related works and opens new perspectives to scientific communities, either for plant biologists interested in using this dataset for their own species or for computer vision experts interested in developing new computer vision tools for plant phenotyping.

Appendix

The appendix contains 5 tables listing the trials. They describe the cameras, the stages used and the number of images. See Tables 6, 7, 8, 9, and 10.

Table 6 Synthesis of trial information - part 1

ID	Plants	Start	End	Cameras	Images	Trays	Pots
1	Rapeseed	14/01/2022	28/01/2022	Intel4	1317	Tray1	40
						Tray2	84
				Intel7	1302	Tray1	40
						Tray2	84
2	Tomatoes	15/03/2022	29/03/2022	Intel4	1310	Tray1	40
						Tray2	40
				Intel5	1259	Tray1	40
						Tray2	40
				Intel6	1310	Tray1	40
						Tray2	40
				Intel7	1312	Tray1	40
						Tray2	40
				Intel9	1268	Tray1	40
						Tray2	40
				Intel10	1270	Tray1	40
						Tray2	40
3	Beans	29/03/2022	08/04/2022	Intel4	965	Tray1	40
						Tray2	40
				Intel5	967	Tray1	40
						Tray2	40
				Intel6	890	Tray1	40
						Tray2	40
				Intel7	714	Tray1	40
						Tray2	40
				Intel9	958	Tray1	40
						Tray2	40
				Intel10	799	Tray1	40
						Tray2	40

Acquisitions 1–3

Table 7 Synthesis of trial information - part 2

ID	Start	End	Plants	Cameras	Trays	Pots	Images
4	Beans	12/04/2022	26/04/2022	Intel6	1329	Tray1	40
				Intel9	1130	Tray1	40
5	17/05/2022	24/05/2022	Beans	Intel4	Tray1	40	655
					Tray2	40	
				Intel5	Tray1	40	644
					Tray2	40	
				Intel6	Tray1	40	646
					Tray2	40	
				Intel7	Tray1	40	653
					Tray2	40	
				Intel9	Tray1	40	652
					Tray2	40	
				Intel10	Tray1	40	518
					Tray2	40	
6	02/09/2022	16/09/2022	Beans	Intel4	Tray1	40	655
					Tray2	40	
				Intel5	Tray1	40	644
					Tray2	40	
				Intel6	Tray1	40	646
					Tray2	40	
				Intel7	Tray1	40	653
					Tray2	40	
				Intel9	Tray1	40	652
					Tray2	40	
				Intel10	Tray1	40	518
					Tray2	40	

Acquisitions 4–6

Table 8 Synthesis of trial information - part 3

ID	Start	End	Plants	Cameras	Trays	Pots	Images
7	16/09/2022	29/09/2022	Rapeseed	Intel4	Tray1	84	1231
					Tray2	84	
				Intel9	Tray1	84	1240
					Tray2	84	
	22/09/2022	06/10/2022		Intel5	Tray1	84	1335
					Tray2	84	
				Intel6	Tray1	84	1092
				Intel8	Tray1	84	1273
					Tray2	84	
	29/09/2022	13/10/2022		Intel4	Tray1	84	1299
					Tray2	84	
				Intel9	Tray1	84	1212
					Tray2	84	
	06/10/2022	20/10/2022		Intel5	Tray1	84	1263
					Tray2	84	
				Intel6	Tray1	84	1333
					Tray2	84	
				Intel8	Tray1	84	1321
8	20/10/2022	31/10/2022	Tomatoes	Intel3	Tray1	40	484
				Intel4	Tray1	40	1029
					Tray2	40	
				Intel6	Tray1	40	1029
					Tray2	40	
				Intel6	Tray1	40	1004
	26/10/2022	08/11/2022		Intel7	Tray1	40	1225
				Intel8	Tray1	40	1130
					Tray2	40	
				Intel9	Tray1	40	1231
					Tray2	40	
				Intel10	Tray1	40	1151
					Tray2	40	

Acquisitions 7 and 8

Table 9 Synthesis of trial information - part 4

ID	Start	End	Plants	Cameras	Trays	Pots	Images
9	08/11/2022	22/11/2022	Tomatoes	Intel3	Tray1	40	1122
				Intel4	Tray1	40	1292
					Tray2	40	
				Intel5	Tray1	40	1311
					Tray2	40	
				Intel6	Tray1	40	1207
				Intel7	Tray1	40	1281
				Intel8	Tray1	40	1310
					Tray2	40	
				Intel9	Tray1	40	1310
					Tray2	40	
				Intel10	Tray1	40	1223
10	01/12/2022	08/12/2022	Beans	Intel3	Tray1	40	645
					Tray2	40	
				Intel4	Tray1	40	645
					Tray2	40	
				Intel5	Tray1	40	646
					Tray2	40	
				Intel6	Tray1	40	644
					Tray2	40	
				Intel7	Tray1	40	644
					Tray2	40	
				Intel8	Tray1	40	644
					Tray2	40	
				Intel9	Tray1	40	645
					Tray2	40	
				Intel10	Tray1	40	646
					Tray2	40	

Acquisitions 9 and 10

Table 10 Synthesis of trial information - part 5

ID	Start	End	Plants	Cameras	Trays	Pots	Images
11	13/12/2022	20/12/2022	Beans	Intel3	Tray1	40	644
					Tray2	40	
				Intel4	Tray1	40	645
					Tray2	40	
				Intel5	Tray1	40	645
					Tray2	40	
				Intel6	Tray1	40	643
					Tray2	40	
				Intel7	Tray1	40	254
					Tray2	40	
				Intel8	Tray1	40	645
					Tray2	40	
				Intel9	Tray1	40	646
					Tray2	40	
				Intel10	Tray1	40	644
					Tray2	40	

Acquisition 11

Acknowledgements
The work was operated in the PHENOTIC platform node of the french national infrastructure on phenotyping PHENOME and the ANR PHENOME 11-INBS-0012 program. This work was partly supported by the SUCSEED project (ANR-20-PCPA-0009). Félix Mercier gratefully acknowledge la Région des Pays de la Loire for funding his doctoral fellowship.

Author contributions
FM, GC, MB, MM, AS, DR conceived and designed this work. FM, GC, AS, MM, MB carried out the acquisitions. FM, GC, managed the data storage. FM, GC, AS, MM, MB carried out image annotations. FM, GC, DR conceived and designed all the data. FM performed the deep learning experiment. FM, AEG, NB, DR interpreted all the data. FM, GC, AEG, NB wrote and revised the manuscript. AEG, NB, MB, DR supervised the work. All authors read and approved the final manuscript.

Funding
PPPR SUCSEED under the management of ANR 20-PCPA-0009.

Availability of data and materials
The datasets generated and/or analyzed during the current study are available in the DATA INRAE repository, DOI: https://doi.org/10.57745/AMFJTK.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no Competing interests.

Received: 22 August 2024 Accepted: 26 January 2025
Published online: 11 February 2025

References

1. Finch-Savage WE, Bassel GW. Seed vigour and crop establishment: extending performance beyond adaptation. J Exp Bot. 2016;67(3):567–91.
2. Wagner M, Demilly D, Ducournau S, Dürr C, Léchappé J. Computer vision for monitoring seed germination from dry state to young seedlings. Seed Test. 2011;142:49–51.
3. Samiei S, Rasti P, Ly VuJ, Buitink J, Rousseau D. Deep learning-based detection of seedling development. Plant Methods. 2020;16:1–11.
4. Garbouge H, Rasti P, Rousseau D. Enhancing the Tracking of Seedling Growth Using RGB-Depth Fusion and Deep Learning. Sensors. 2021;21(24):8425.
5. Couasnet G, Cordier M, Garbouge H, Mercier F, Pierre D, Ghaziri AE, et al. Growth Data-An automatic solution for seedling growth analysis via RGB-Depth imaging sensors. SoftwareX. 2023;24:101572. https://doi.org/10.1016/j.softx.2023.101572.
6. Intel. Intel, editor.: Product Family D400 Series Datasheet. Intel. Rev. 018. https://www.intelrealsense.com/.
7. Kurtser P, Lowry S. RGB-D datasets for robotic perception in site-specific agricultural operations-A survey. Comput Electron Agric. 2023;212:108035. https://doi.org/10.1016/j.compag.2023.108035.
8. Das Choudhury S, Samal A, Awada T. Leveraging image analysis for high-throughput plant phenotyping. Front Plant Sci. 2019. https://doi.org/10.3389/fpls.2019.00508.
9. Lu Y, Young S. A survey of public datasets for computer vision tasks in precision agriculture. Comput Electron Agric. 2020;178:105760. https://doi.org/10.1016/j.compag.2020.105760.
10. Cao Y, Tian D, Tang Z, Liu X, Hu W, Zhang Z, et al. OPIA: an open archive of plant images and related phenotypic traits. Nucleic Acids Res. 2024;52(D1):D1530–7.
11. Cruz JA, Yin X, Liu X, Imran SM, Morris DD, Kramer DM, et al. Multi-modality imagery database for plant phenotyping. Mach Vis Appl. 2016;27(5):735–49. https://doi.org/10.1007/s00138-015-0734-6.
12. Scharr H, Minervini M, Fischbach A, Tsaftaris SA. Annotated Image Datasets of Rosette Plants. 2014. https://juser.fz-juelich.de/record/154525.
13. Arrechea-Castillo DA, Espitia-Buitrago P, Arboleda RD, Hernandez LM, Jauregui RN, Cardoso JA. High-resolution image dataset for the automatic classification of phenological stage and identification of racemes in Urochloa spp. hybrids. Data Brief. 2024;57:110928.

14. Uchiyama H, Sakurai S, Mishima M, Arita D, Okayasu T, Shimada A, et al. An Easy-To-Setup 3D Phenotyping Platform for KOMATSUNA Dataset. In: Proceedings of the IEEE International Conference on Computer Vision (ICCV) Workshops; 2017; 5.
15. Gené-Mola J, Vilaplana V, Rosell-Polo JR, Morros JR, Ruiz-Hidalgo J, Gregorio E. KFuji RGB-DS database: Fuji apple multi-modal images for fruit detection with color, depth and range-corrected IR data. Data Brief. 2019;25:104289. https://doi.org/10.1016/j.dib.2019.104289.
16. Lac L, Keresztes B, Louargant M, Donias M, Da Costa JP. An annotated image dataset of vegetable crops at an early stage of growth for proximal sensing applications. Data Brief. 2022;42:108035. https://doi.org/10.1016/j.dib.2022.108035.
17. Genze N, Vahl WK, Groth J, Wirth M, Grieb M, Grimm DG. Manually annotated and curated dataset of diverse weed species in maize and sorghum for computer vision. Sci Data. 2024;11(1):109. https://doi.org/10.1038/s41597-024-02945-6.
18. Liang Z, Pandey P, Stoerger V, Xu Y, Qiu Y, Ge Y, et al. Conventional and hyperspectral time-series imaging of maize lines widely used in field trials. GigaScience. 2017;7(2):gix117. https://doi.org/10.1093/gigascience/gix117.
19. Arend D, Lange M, Pape JM, Weigelt-Fischer K, Arana-Ceballos F, Mücke I, et al. Quantitative monitoring of Arabidopsis thaliana growth and development using high-throughput plant phenotyping. Sci Data. 2016;3(1):160055. https://doi.org/10.1038/sdata.2016.55.
20. Arrechea-Castillo DA, Espitia-Buitrago P, Arboleda RD, Hernandez LM, Jauregui RN, Cardoso JA. High-resolution image dataset for the automatic classification of phenological stage and identification of racemes in Urochloa spp. hybrids. Data Brief. 2024;57:110928. https://doi.org/10.1016/j.dib.2024.110928.
21. Demirkaya A, Chen J, Oymak S. Exploring the Role of Loss Functions in Multiclass Classification. 2020 54th Annual Conference on Information Sciences and Systems (CISS). 2020; 1–5.
22. Kingma DP, Ba J. Adam: a method for stochastic optimization. CoRR. 2014; abs/1412.6980.
23. Garbouge H, Rasti P, Rousseau D. Deep learning-based detection of seedling development from indoor to outdoor. In: International conference on systems, signals and image processing. Springer. 2021; 121–131.
24. Cao W, Mirjalili V, Raschka S. Rank consistent ordinal regression for neural networks with application to age estimation. Pattern Recognit Lett. 2020;140:325–31. https://doi.org/10.1016/j.patrec.2020.11.008.
25. Shi X, Cao W, Raschka S. Deep neural networks for rank-consistent ordinal regression based on conditional probabilities. Pattern Anal Appl. 2023;26(3):941–55. https://doi.org/10.1007/s10044-023-01181-9.
26. Kumar S, Luo W, Kantor G, Sycara K. Active learning with gaussian processes for high throughput phenotyping. arXiv preprint arXiv:1901.06803. 2019.
27. Mercier F, Bouhlel N, El Ghaziri A, Online Rousseau D, Sampling Bayesian Adaptive, for Nonlinear Model: Application to Plant Phenotyping. In,. 32nd European Signal Processing Conference (EUSIPCO). IEEE. 2024; 2024: 2537–41.
28. Gómez-Carmona O, Casado-Mansilla D, Kraemer FA, López-de Ipiña D, García-Zubia J. Exploring the computational cost of machine learning at the edge for human-centric internet of things. Future Gener Comput Syst. 2020;112:670–83.
29. Tekin N, Aris A, Acar A, Uluagac S, Gungor VC. A review of on-device machine learning for IoT: an energy perspective. Ad Hoc Netw. 2024;153:103348.
30. Morris D. A Pyramid CNN for Dense-Leaves Segmentation. In: 2018 15th Conference on Computer and Robot Vision (CRV); 2018; 238–245.
31. Wang M, Fu B, Fan J, Wang Y, Zhang L, Xia C. Sweet potato leaf detection in a natural scene based on faster R-CNN with a visual attention mechanism and DIoU-NMS. Ecol Inform. 2023. https://doi.org/10.1016/j.ecoinf.2022.101931.
32. Cordier M, Torres C, Rasti P, Rousseau D. On the use of circadian cycles to monitor individual young plants. Remote Sens. 2023. https://doi.org/10.3390/rs15112704.
33. Elsken T, Metzen JH, Hutter F. Neural architecture search: a survey. J Mach Learn Res. 2019;20(55):1–21.
34. Wang F, Zhang M, Wang X, Ma X, Liu J. Deep learning for edge computing applications: a state-of-the-art survey. IEEE Access. 2020;8:58322–36.
35. Yu R, Liu S, Wang X. Dataset distillation: a comprehensive review. IEEE Trans Pattern Anal Mach Intell. 2023. https://doi.org/10.1109/TPAMI.2023.332337.
36. Mavaie P, Holder L, Skinner MK. Hybrid deep learning approach to improve classification of low-volume high-dimensional data. BMC Bioinform. 2023;24(1):419.
37. Procopio A, Cesarelli G, Donisi L, Merola A, Amato F, Cosentino C. Combined mechanistic modeling and machine-learning approaches in systems biology-a systematic literature review. Comput Methods Progr Biomed. 2023;240:107681.
38. Pang Y, Lin J, Qin T, Chen Z. Image-to-image translation: methods and applications. IEEE Trans Multimedia. 2021;24:3859–81.
39. Zhou Y, Hu B, Yuan X, Huang K, Yi Z, Yen GG. Multi-objective evolutionary generative adversarial network compression for image translation. IEEE Trans Evolut Comput. 2023. https://doi.org/10.1109/TEVC.2023.3261135.

Publisher's Note