The Oil Radish Growth Dataset for Semantic Segmentation and Yield Estimation

Anders Krogh Mortensen

Department of Agroecology, Aarhus University
Forsøgsvej 1, 4200 Slagelse, Denmark

anmo@agro.au.dk

Søren Skovsen

Department of Engineering, Aarhus University
Finlandsgade 22, 8210 Aarhus N, Denmark

ssk@eng.au.dk

Henrik Karstoft

Department of Engineering, Aarhus University
Finlandsgade 22, 8210 Aarhus N, Denmark

hka@eng.au.dk

René Gislum

Department of Agroecology, Aarhus University
Forsøgsvej 1, 4200 Slagelse, Denmark

rg@agro.au.dk

Abstract

Data sharing in research is important in order to re-produce results, develop global models, and benchmark methods. This paper presents a dataset containing image and field data from a field plot experiment with oil radish (Raphanus sativus L. var oleiformis) as catch crop after spring barley. The field data consists of fresh weight, dry weight, Carbon content and Nitrogen content from multiple weekly plant samples collected from the plots. The image data consists of images collected weekly prior to the plant samples. A subset of the images corresponding to the plant sampling areas have been annotated pixelwise. In addition to the image and field data, weather data from the growing period is also included in the dataset.

The dataset is accompanied by two challenges: 1) semantic segmentation of crops and 2) oil radish yield esti-mation. The former challenge focuses on data image, while the latter focuses on the field data. Baseline methods and results are provided for both challenges.

1. Introduction

With a world population expecting to reach 10 billion by 2050 [17], there is a need for an increase in agricultural production. With an increasingly environmental concern on especially Nitrogen (N) leaching, it is broadly agreed upon an increase in production must not cause an concomitant negative environmental impact. One possible solution could be to increase production through more knowledge per area [3] or more precisely through sustainable intensification of agriculture. To acquire this knowledge, data and knowledge must be collected from the field, which in turn can be used to build decision support systems for farmers. More knowl-edge per area can be collected by various types of sensors,

e.g. as RGB images which has the advantage of resembling one of the human senses, which make it easier to determine the ground truth. The collected image data must also be ac-companied by field measurements, when the initial model is developed.

The image and in particularly field data collection is very time- and labour consuming and therefore expensive to col-lect. Sharing collected data is therefore obvious not only from the before mentioned reasons, but also due to the fact that more data from different locations, years and crops is necessary in order to develop more global models. It, furthermore, helps reproducing results and benchmarking methods.

1.1. Related work

In recent years, several datasets containing images of plants have been made public available. These datasets may be grouped based on both how plants were grown and how controlled the data collection was performed.

In the fully controlled setting, the growth of the plants is controlled closely and the plants are often grown individu-ally. During image acquisition, the light conditions is also controlled carefully. Examples of these include the 'Plant Phenotyping Datasets' [12] and the 'Plant Seedling Dataset' [8]. The 'Plant Phenotyping Datasets' consist of images and annotations of Arabidopsis and tobacco over several weeks. The annotations include segmentation masks and bound-ing boxes of both plants and individual leafs. The 'Plant Seedling Dataset' consists of images of 12 crop and weed species collected every 2 to 3 days for a 3 week period after germination.

In the semi-controlled setting, the plants are grown out-side and subject to the weather. During data collection, however, special attention is paid towards the light condi-tions through shielding the sun or using a powerful flash.

Datasets collected in semi-controlled settings include the 'Sugar Beet Dataset' [4], 'The Crop/Weed Field Image Dataset' [9] and the 'Spatial and Temporal Clovergrass Dataset' [1]. The 'Sugar Beet Dataset' consists of RGB images, NIR images, lidar and laser scans of sugar beets collected 2 to 3 times per week over a 3 month growing period using a field robot. The annotations consists of 300 images labelled according to non-vegetation, sugar beet and 9 types of weed species[1]. The 'Crop/Weed Field Image Dataset' consists of images of carrot and weed collected using a field robot. The images are accompanied by automatically generated vegetation masks, which were subsequently manually annotated as weed or crop. The 'Spatial and Temporal Clovergrass Dataset' consists of images and botanical measurements collected from clovergrass leys at two locations. The images are labelled according to clover, grass, weed and soil, but the labels cannot be used for training, as they are generated using image analysis [2].

In the uncontrolled setting, the plants are grown outside, but in contrast to the semi-controlled case, no special attention is made to the light conditions taken when collecting the image data. The 'Leaf Counting Dataset' from [16] is an example of such a dataset, where the images have been collected by several people under various conditions using smartphones, consumer cameras and industrial cameras. The dataset consists cropped images of weeds, which have been annotated with the number of leafs present connect to each plant.

As an alternative to manually annotate a large dataset, [6] and [15] have created synthetic images of weeds and clovergrass, respectively. The synthetic images are created by cropping plants from real images and placing the randomly on a soil background. This allows them to create large image datasets with complete pixel-wise annotations from a small number of samples.

A common trait for the majority of the above mentioned datasets that they either do not include any field data (e.g. biomass) or the link between the annotated image data and field data is weak (e.g. 'Spatial and Temporal Clovergrass Dataset'). This paper presents the "Oil Radish Growth Dataset" consisting of images and field data with a strong link between the two. The dataset was collected over a two month period in an uncontrolled outdoor setting. It contains pixelwise annotated images of oil radish and volunteer seeded barley with corresponding fresh weight, dry weight, Carbon (C) content and N content of oil radish, barley, weed and stubble. The dataset is accompanied by two challenges (including baseline results): 1) semantic segmentation and 2) yield estimation.

This paper is structured as follows. Section 2 describes how the image and field data were collected. Section 3 describes how the data forms the dataset, and Section 4

presents two challenges with baseline results. The dataset is discussed and concluded in section 5.

2. Field experiment and data acquisition

The image and field data included in the dataset is based on the work of [14]. Besides making the dataset openly available, the present paper greatly increases the number of labelled images compated to [14]. The following subsections describe the field experiment and the data collection.

2.1. Field experiment

The field plot experiment was established in late April 2015 near Foulum, Denmark by sowing spring barley (*Hordeum vulgare* L.). On 11th of August, the spring barley was harvested. On 20th of August, 36 beds (each 15 m × 3 m) were established in the harvested field. Barley stubble were left in the field. The beds were established to evaluate the effect of nine different N catch crop strategies with four repetitions of each strategy. In the scope of this paper, only the four beds using oil radish (*Raphanus sativus* L. var oleiformis) as a catch crop is considered.

To ensure an approximately uniform sampling from the beds, sampling areas (0.5 m × 0.5 m) were pre-allocated by dividing each bed into six regions with eight sampling areas each except for the end regions. The sampling areas were located at the edge of the beds to reduce the effect on another experiment. A yellow stick was placed at the corner of each sampling area prior to the first week of data collection. The yellow stick helped locate the sampling areas when collecting field data in subsequent weeks and subsequently in the images.

Each week, one sampling area from each of the regions was selected for biomass collection. The selected sampling areas were selected to avoid sampling a neighbouring sampling area in the subsequent week.

From late September till early November, weekly image and plant samples were collected from the beds with oil radish. The weekly data collection was carried out by first scanning each bed with an RGB camera and subsequently collecting plant samples from the sampling area. Plant samples were collected for seven weeks, and images were collected for eight weeks. The images collected in a given week were used to locate the exact sampling areas in the images from the previous week. Hence, the image collection continued one more week than the plant sampling.

2.2. Image data acquisition

The beds were scanned using Sony α7 24 Mpx full frame RGB camera (6000 px × 4000 px) with a 35 mm lens. The camera was mounted downward facing in front of tractor. The camera height (±20 cm) and angle (±5°) varied slightly between weeks. The cameras field of view covered

---

[1]The annotations have been expanded upon after publication.

Figure 1. Camera setup used for image acquisition.

approximately 2 m × 1.4 m at ground level and a ground sampling distance of 0.33 mm px^{-1}. The camera was offset to the right such that the field of view captured mainly the right side of the plot, when driving through it (Figure 1).

A real-time kinematic (RTK) GPS was positioned directly above the camera to log the position of each image. The camera and GPS were connected to a laptop located in the cabin of the tractor. Starting and stopping image capturing were performed manually through the laptop before entering and after exiting each plot, respectively. Captured images were transferred to the laptop through USB. The GPS position was stored every 1 s and used to geotag the images offline.

Each plot was traversed once in each direction each week in order to cover the entire plot. The plots were traversed at approximately 0.05 m s^{-1} to maintain a large overlap between the images.

2.3. Field data acquisition

Plant samples were collected by placing a 0.5 m × 0.5 m cutting frame at the sampling area and then collecting all above-ground plant material including stubble inside the cutting frame. Plant material originating inside the cutting frame but ending outside was partitioned where it naturally crossed the imaginary vertical border of the cutting frame. The same procedure was applied to plant material originating outside the cutting frame. This was done to achieve a better correlation with what was observed in the images. The plant samples were collected from the sampling areas on the same day as the images were collected. In total, 136 plant samples were collected during the seven weeks of field data collection.

The plant samples were stored at 4 °C for 0 to 7 days before they were analyzed. The storing period was determined by the availability of the ovens used in the subsequent drying process. Before weighing the samples, each sample was divided into four fractions (oil radish, barley/grass, weed

and stubble), which were subsequently analyzed individually. Each fraction-sample was weighed to determine the fresh weight, dried and then weighed again to determine the dry weight. The samples were dried for 48 hours at 60 °C to ensure all water had evaporated. After the dry matter analysis, the samples were stored at dry storage until they could be analyzed for N- and C-content. Each fraction-sample was ground and a 5 mg sample was extracted to analyze for N- and C-content. The analysis followed Dumas method using a Vario El III (Elementar Analysensysteme GmbH, Langenselbold, Germany).

3. Dataset

The data collected from the field plot experiment have been organized into the "Oil radish growth dataset" [2]. The dataset consists of the image and field data described in the previous section as well as weather data from a local weather station. The weather data includes daily mean air temperature at 2 m, daily mean soil temperature at 0.10 m, daily mean precipitation, and daily total global radiation from the spring barley harvest date to the last date of data collection (Figure 2).

To create a link between the image data and the field data, the sampling areas of a given week were identified in the images by comparing them to the images of a subsequent week. 10 sampling areas could not be located in the images due missing images due technical issues during data collection. The issue was not discovered until after the field data had been collected, thus, it was not possible to acquire new images of the affected sampling areas. For each sampling area and image pair, a 1600 px × 1600 px patch around the sampling area was extracted from the image. The pixel coordinates of the corners of the sampling area was recorded. Each image of the sampling areas was pixel-wise annotated as either oil radish, barley/grass, weed, soil, equipment, (barley) stubble, or unknown. Grass was included in the barley category to match the fractions used in collected field data. Weed accounted for all green plant material, which did not fit into the oil radish or barley/grass categories. The annotation was performed in two steps. First edges were detected to create blobs. Although, the were detected automatically, the edges could be corrected by the annotator, if they were missing or misplaced. In the second step, the blobs were labelled as one of the previously mentioned classes. In practice, the annotation was performed as an iterative process between the two steps. The annotation was performed by an external company with expertise in image annotation, but not in crops.

Figure 4 illustrates the labelled data (image, field and weather data) from plot 3 in week 5.
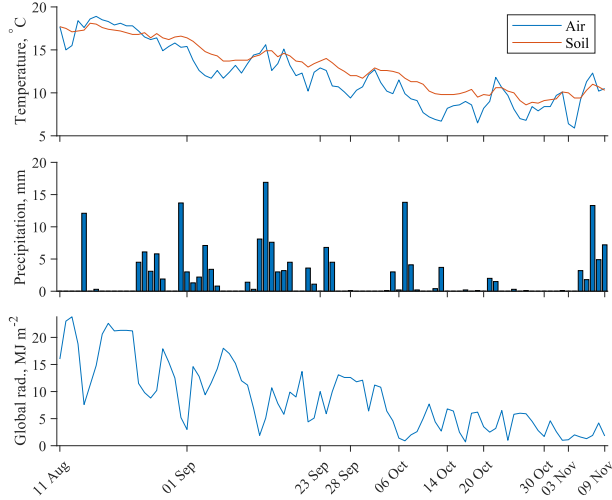

Figure 2. Weather data included in the dataset ranging from spring barley harvest date to the last date of data collection. **Top**: Daily average air and soil temperature at 2 m height and 10 cm depth, respectively. **Middle**: Daily total precipitation. **Bottom**: Global radiation.

	Training set	Test set
Sampling areas	101	35
Labelled images	95	34
Unlabelled images	5287	–

Table 1. Summary of sampling areas and image data in the training set and test set.

A large set of images was collected during the field plot experiment, but only a small portion of the images was labelled. The images, which was not associated with a sampling area, have been included as unlabelled images to facilitate unsupervised approaches.

The dataset was split into a training set and a test set (Table 1) based on which plots, the data was collected from. Plot 4 was selected for the test set, while plot 1 to 3 was selected for training set.

The image and field data from the training set is summarized in Figure 3.

4. Challenges

The presented dataset is accompanied by two challenges: 1) Semantic segmentation and 2) Oil radish yield estimation. The challenges are available as a permanent competition online[3].

4.1. Semantic segmentation

The end goal of this dataset is to estimate the yield. In order to include the image data, information regarding the yield must be extracted from the image data. To do so, a
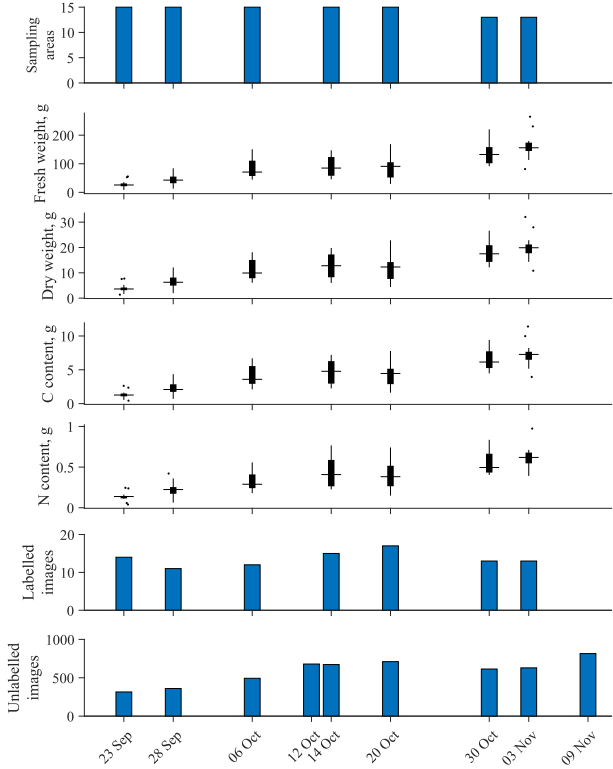
Figure 3. Summary of oil radish field data and image data from the training set grouped by date of collection. From top to bottom: 1) Barplot of number of sampling areas. 2-5) Boxplot of oil radish fresh weight, dry weight, C-content and N-content. Horizontal lines indicate median. Edges of boxes indicate 25th and 75th percentile values. End of vertical lines indicate most extreme non-outlier data points. · indicate outliers. 6) Barplot of number of annotated images. 7) Boxplot of number of unlabelled images. Note the two extra dates with unlabelled images.

first potential step is to determine which pixels belongs to which species.

As such, in the semantic segmentation challenge, the task is to classify each pixel in the labelled images from the training set. Each pixel must be classified as either oil radish, barley/grass, weed, soil, equipment or (barley) stubble. Pixels labelled as unknown in manual annotations are ignored during evaluation and can thus be ignored during training. Pixels classified as unknown are however not ignored and will potentially count negatively towards the final score.

4.1.1 Evaluation metrics

The semantic segmentation challenge is evaluated using the intersection over union (or Jaccard index [10]), which is commonly used when evaluating semantic segmentation tasks [7, 11].

Let $C = [c_{o,p}]$ be a confusion matrix, where $c_{o,p}$ is the

Image data

Sampling areas

Annotations
- ☐ Unknown
- 🟥 Oil radish
- 🟩 Barley
- 🟪 Weed
- 🟨 Stubble
- 🟦 Soil
- 🟩 Equipment

Field data

Fresh weight:					
Radish	72.5 g	95.7 g	41.6 g	91.9 g	29.5 g
Barley	19.6 g	16.3 g	33.6 g	9.4 g	23.0 g
Weed	0.6 g	1.5 g	2.9 g	6.0 g	9.6 g
Stubble	17.2 g	13.2 g	13.9 g	9.1 g	17.0 g
Dry weight					
Radish	10.2 g	12.9 g	5.9 g	12.3 g	4.4 g
Barley	2.8 g	3.8 g	8.9 g	2.5 g	7.0 g
Weed	0.3 g	0.3 g	0.8 g	1.6 g	2.0 g
Stubble	13.7 g	10.9 g	10.8 g	7.9 g	13.5 g
C-content					
Radish	3.66 g	4.59 g	2.14 g	4.50 g	1.62 g
Barley	0.97 g	1.40 g	3.61 g	0.90 g	2.79 g
Weed	0.10 g	0.11 g	0.29 g	0.57 g	0.77 g
Stubble	4.44 g	3.63 g	4.71 g	3.13 g	5.25 g
N-content					
Radish	0.36 g	0.44 g	0.20 g	0.44 g	0.15 g
Barley	0.06 g	0.10 g	0.19 g	0.06 g	0.14 g
Weed	0.00 g	0.01 g	0.02 g	0.04 g	0.05 g
Stubble	0.10 g	0.09 g	0.06 g	0.05 g	0.01 g

Weather data

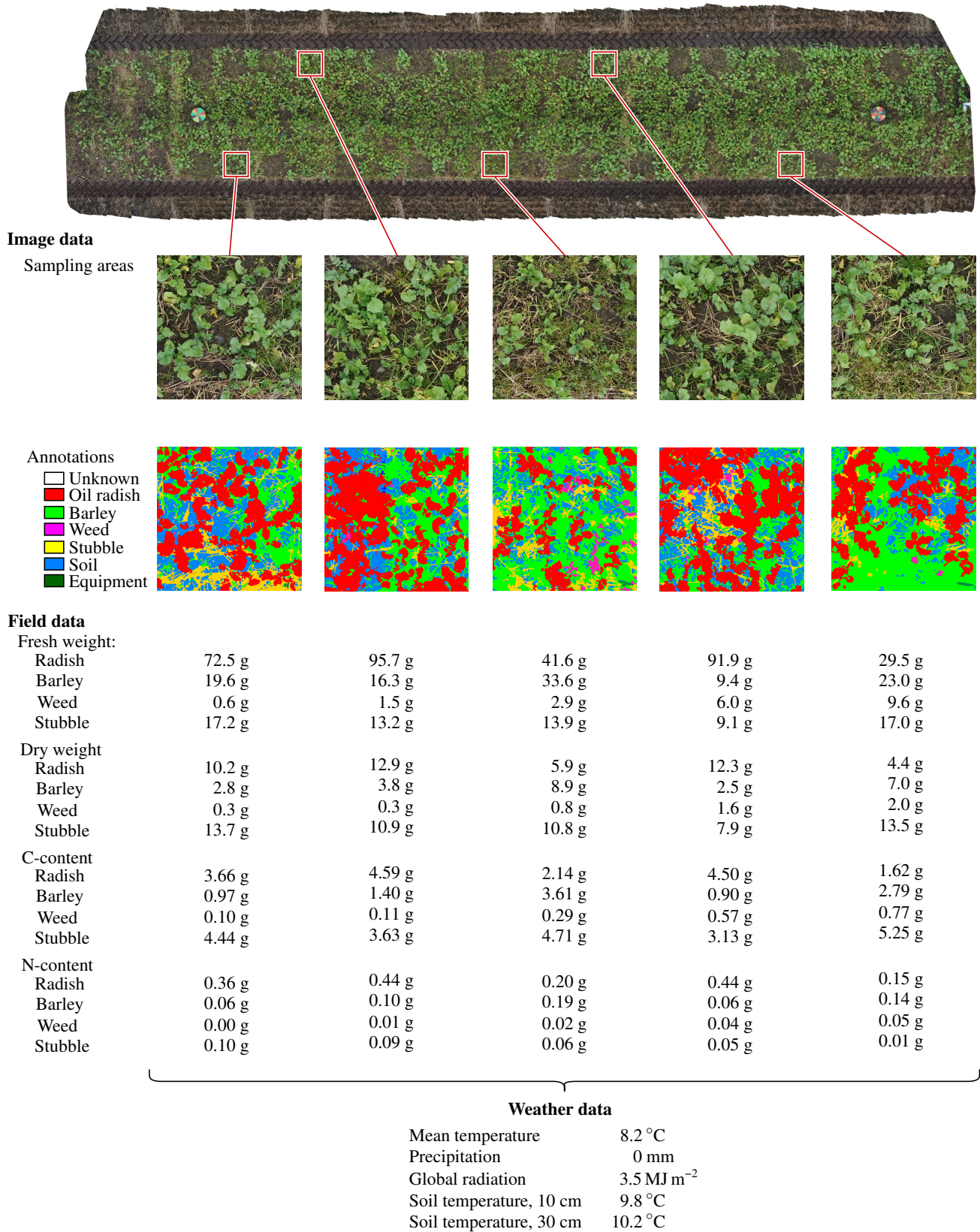| Mean temperature	8.2 °C
Precipitation	0 mm
Global radiation	3.5 MJ m^{-2}
Soil temperature, 10 cm	9.8 °C
Soil temperature, 30 cm	10.2 °C

Figure 4. Example of available data from the dataset. The example shows plot 3 in week 5 with the associated sampling areas, annotated images, field data and weather data.

number of instances from class o predicted as class p. If $o = p$, $c_{o,p}$ is the number of correct predictions for a given class. The intersection over union for class k is then given by:

$$IoU_k = \frac{c_{k,k}}{-c_{k,k} + \sum_{o=1}^{K} c_{o,k} + \sum_{p=1}^{K} c_{k,p}} \quad (1)$$

where K is the number of classes. The per class, mean and frequency weighted intersection over union are used as evaluation metrics. The mean and frequency weighted intersection over union are given by:

$$mean\ IoU = \frac{1}{K} \sum_{k=1}^{K} IoU_k \quad (2)$$

and

$$f.w.\ IoU = \sum_{k=1}^{K} \left(IoU_k \frac{\sum_{o=1}^{K} c_{o,k}}{\sum_{o=1}^{K} \sum_{p=1}^{K} c_{o,p}} \right) \quad (3)$$

Pixels labelled as 'Unknown' in the annotated images are ignored in the final evaluation.

4.1.2 Baseline results

Two deep convolutional neural networks were trained to provide baseline models for this challenge. Both networks used the FCN-8s network structure [11], but one network was trained solely on the labelled images, while the other was trained using the unlabelled images. Both networks was trained for 50 epochs with a learning rate of 25×10^{-5} using the Adagrad algorithm for optimization [5] and a batch size of 3. Each class was weighed with respect to the log frequency of its occurrence in the training set. The class weights was normalized such that the average class weight was 1. During training, the images were augmented using random brightness, saturation, hue and contrast and randomly cropping the image to $1200\,\mathrm{px} \times 1200\,\mathrm{px}$. 8% of the image data (8 labelled images and 423 unlabelled images) was used for validation to monitor the performance during training.

The network trained on the labelled data was trained first. To train FCN-8s on the unlabelled data, the network trained on the labelled data was used to 'annotate' the unlabelled data following the example of [13]. Only the center $3000\,\mathrm{px} \times 2000\,\mathrm{px}$ from the unlabelled images were used.

The baseline results for both networks are summarized in Table 2.

4.2. Oil radish yield estimation

The oil radish yield estimation challenge focuses on the collected oil radish field data. In the field plot experiment, oil radish was sown as a catch crop and used to capture N in soil in order to minimize N-leaching to ground and surface water. The captured N will be available for the crops

sown in the following spring and may be included in the N accounting for crop rotation the following year.

The challenge is, therefore, to predict the field data associated with oil radish (fresh weight, dry weight, C-content and N-content) from the sampling areas. Though the challenge stems from the field data, it is encouraged to include the image and weather data in the predictions. Likewise, it is encouraged to include field data from the other fractions (barley/grass, weed and stubble) to account for any potential interactions.

4.2.1 Evaluation metrics

Yield estimation models are evaluated using the root mean square error and the mean absolute percentage error. Given a set of observed values $\{t_1, t_2, \ldots, t_N\}$ and a set of predicted values $\{y_1, y_2, \ldots, y_N\}$, the root mean square error (RMSE) and mean absolute percentage error (MAPE) is given by:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (t_i - y_i)^2} \quad (4)$$

and

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} \left| \frac{t_i - y_i}{t_i} \right| 100\% \quad (5)$$

where t_i and y_i are the ith observed and predicted values, respectively, and N is the total number of samples. During evaluation, sampling areas without any corresponding image data are ignored.

4.2.2 Baseline results

Linear models for predicting oil radish dry weight, fresh weight, C-content and N-content have been fitted to the training data to create baselines for this challenge. Two sets of models have been fitted.

The first set of models are based on the temperature sum derived from the weather data. The temperature sum T_{sum} over a given period is given by:

$$T_{sum} = \sum_{t=t_0}^{t_N} max(T_t - T_B, 0) \quad (6)$$

where T_t is the daily average temperature to the time t, T_B is the base temperature required for growth, and t_0 and t_N is the starting and ending time. For the baseline models, $T_B = 0$ and t_0 was set to the harvest date of spring barley. Scatter plots of the training data and the fitted models are shown in Figure 5 (top row).

The second set of models are based on the oil radish coverage derived from the image data. The coverage of a given

Method	Intersection over union							
	$mean$	$f.w.$	Radish	Barley	Weed	Stubble	Soil	Equipment
FCN-8s	0.449	0.712	0.859	0.459	0.135	0.322	0.633	0.285
FCN-8s + UL	0.449	0.712	0.859	0.461	0.119	0.325	0.631	0.300

Table 2. Baseline results for the semantic segmentation challenge. 'FCN-8s' was trained only on the labelled images. 'FCN-8s + UL' was trained on the unlabelled images. F.w. is the frequency weighted intersection over union.

sampling area was calculated as the number pixels labelled as oil radish with respect to the total number of pixels. For each sampling area, the full 1600 px × 1600 px image was used. For evaluation on the test set, the label predictions produced by the FCN-8s network trained in the semantic segmentation challenges were used. Scatter plots of the training data and the fitted models are shown in Figure 5 (bottom row).

The oil radish yield estimation baseline results are summarized in Table 3. From both the scatter plots of the training data (Figure 5) and the baseline results from the evaluation on the test set (Table 3) it is clear that coverage extracted from the images is a better predictive variable than the temperature sum. This is largely due to the fact that the coverage can explain the spatial variance within the oil radish. The temperature sum, on the other hand, is the same for all sampling areas within the plots one a given date and thus can not be used to explain this variance. The coverage can, however, only be expected to provide decent results up till a certain point. When the oil radish covers the full image, an increase in biomass can no longer be described by the coverage and must be explained otherwise. One potential solution would be to create a joint model based on both image and weather data.

5. Discussion and conclusion

This paper has presented the Oil Radish Growth dataset, which contains image and field data from an oil radish field plot experiment collected over 8 weeks. To the best of the authors knowledge, this is the first openly available dataset which combines images and field data of oil radish.

One drawback of the dataset is that it contains data from only a single growing season and from a single field. This severely limits the potential of creating generalized models based on the dataset. Especially considering years with extreme growth such as 2014, where the conditions for growth in the fall was just right. The dataset will, however, still be useful for benchmarking and may be combined with other future datasets to create larger datasets.

The image annotations was performed by an external company with no expertise in crops. As such, the annotations are expected to contain some level of mislabelling. The mislabelled annotations have been corrected when identified, but some may still inhabit the dataset.

The authors remain open to any feedback, corrections to or extensions on the dataset and aim to update the dataset based on those responses. Possible extensions to the dataset annotations include plant instance labels, stem emergence points and labelling of individual leafs (similar to [12]). Each of which can easily spur off an additional challenge.

References

[1] A. K. Bakken, H. Bonesmo, and B. Pedersen. Spatial and temporal abundance of interacting populations of white clover and grass species as assessed by image analyses. *Dataset Papers in Science*, 2015, 2015. 2

[2] H. Bonesmo*, K. Kaspersen, and A. Kjersti Bakken. Evaluating an image analysis system for mapping white clover pastures. *Acta Agriculturae Scandinavica, Section B-Soil & Plant Science*, 54(2):76–82, 2004. 2

[3] A. Buckwell, A. Heissenhuber, and W. Blum. The sustainable intensification of european agriculture. Technical report, RISE Foundation, 2013. 1

[4] N. Chebrolu, P. Lottes, A. Schaefer, W. Winterhalter, W. Burgard, and C. Stachniss. Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. *The International Journal of Robotics Research*, 36(10):1045–1052, 2017. 2

[5] J. Duchi, E. Hazan, and Y. Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research*, 12(Jul):2121–2159, 2011. 6

[6] M. Dyrmann, A. Mortensen, H. Midtiby, and R. Nyholm Jørgensen. Pixel-wise classification of weeds and crops in images by using a fully convolutional neural network. 2016. Paper No. 608, CIGR 2016; Conference date: 26–29 June 2016. 2

[7] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 111(1):98–136, Jan 2015. 4

[8] T. M. Giselsson, M. Dyrmann, R. N. Jørgensen, P. K. Jensen, and H. S. Midtiby. A Public Image Database for Benchmark of Plant Seedling Classification Algorithms. *arXiv preprint*, 2017. 1

[9] S. Haug and J. Ostermann. A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks. In L. Agapito, M. M. Bronstein, and C. Rother, editors, *Computer Vision - ECCV 2014 Workshops*, pages 105–116. Springer International Publishing, 2015. 2

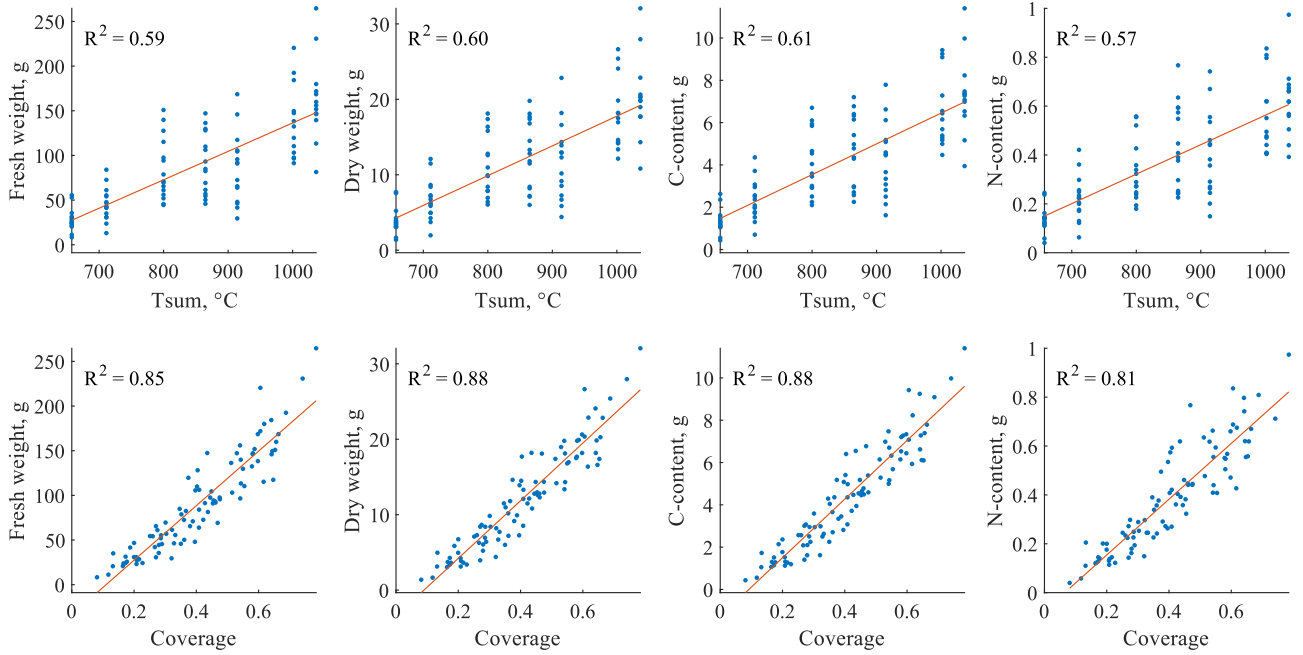[10] P. Jaccard. The distribution of the flora in the alpine zone.1. *New Phytologist*, 11(2):37–50, 1912. 4

Figure 5. Baseline models for radish yield estimation fitted to the training set. **Top**: Models based on temperature sum ($T_B = 0$ °C). **Bottom**: Models based on radish coverage estimated from the annotated images. The dots are the data points and the lines are the fitted models.

Predictor	Fresh weight		Dry weight		C-content		N-content	
	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE
T_{sum}	72.3 g	35.0%	7.5 g	29.6%	2.74 g	29.9%	0.29 g	32.1%
Coverage	36.0 g	17.9%	3.4 g	17.2%	1.24 g	17.0%	0.14 g	15.0%

Table 3. Baseline results for the yield prediction challenge.

[11] J. Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3431–3440, 2015. 4, 6

[12] M. Minervini, A. Fischbach, H. Scharr, and S. A. Tsaftaris. Finely-grained annotated datasets for image-based plant phenotyping. *Pattern Recognition Letters*, 81:80 – 89, 2016. 1, 7

[13] A. Mortensen. *Estimation of Above-Ground Biomass and Nitrogen-Content of Agricultural Field Crops using Computer Vision*. PhD thesis, Department of Agroecology, Aarhus University, 2018. 6

[14] A. Mortensen, M. Dyrmann, H. Karstoft, R. Nyholm Jørgensen, and R. Gislum. Semantic segmentation of mixed crops using deep convolutional neural network. 2016. Paper No. 259, CIGR 2016; Conference date: 26–29 June 2016. 2

[15] S. Skovsen, M. Dyrmann, A. Mortensen, K. Steen, O. Green, J. Eriksen, R. Gislum, R. Nyholm Jørgensen, and H. Karstoft. Estimation of the botanical composition of clover-grass leys from rgb images using data simulation and fully convolutional neural networks. *Sensors*, 17(12), 12 2017. Special Issue: Sensors in Agriculture. 2

[16] N. Teimouri, M. Dyrmann, P. R. Nielsen, S. K. Mathiassen, G. J. Somerville, and R. N. Jørgensen. Weed growth stage estimator using deep convolutional neural networks. *Sensors*, 18(5), 2018. 2

[17] United Nations. World Population Prospects: The 2017 Revision, Key Findings and Advance Tables. Technical report, United Nations, Department of Economic and Social Affairs, Population Division, 2017. Working Paper No. ESA/P/WP/248. 1