A Crop/Weed Field Image Dataset for the Evaluation of Computer Vision Based Precision Agriculture Tasks

Sebastian Haug[1(✉)] and Jörn Ostermann[2]

[1] Corporate Research, Robert Bosch GmbH, Stuttgart, Germany
sebastian.haug@de.bosch.com
[2] Leibniz Universität Hannover, Hannover, Germany
ostermann@tnt.uni-hannover.de

Abstract. In this paper we propose a benchmark dataset for crop/weed discrimination, single plant phenotyping and other open computer vision tasks in precision agriculture. The dataset comprises 60 images with annotations and is available online (http://github.com/cwfid). All images were acquired with the autonomous field robot Bonirob in an organic carrot farm while the carrot plants were in early true leaf growth stage. Intra- and inter-row weeds were present, weed and crop were approximately of the same size and grew close together. For every dataset image we supply a ground truth vegetation segmentation mask and manual annotation of the plant type (crop vs. weed). We provide initial results for the phenotyping problem of crop/weed classification and propose evaluation methods to allow comparison of different approaches. By opening this dataset to the community we want to stimulate research in this area where the current lack of public datasets is one of the barriers for progress.

Keywords: Computer vision · Phenotyping · Dataset · Precision agriculture · Classification · Bonirob field robot

1 Introduction

Automation in agriculture, intelligent farm management as well as robotic precision agriculture activities require detailed information about the environment, the field, the condition and the phenotype of individual plants. An increase in available data allows more automatic, precise, cost-effective and organic production of crops and vegetables.

Camera sensors and computer vision with machine learning are promising technologies to capture such information and further process it to be able to realize autonomous farming. Combined with field robots such as Bonirob [18] that navigate autonomously in fields [1,9] tasks that are still manual today can be automated. For example, weed control in organic carrot farming is still performed manually and necessary to avoid substantial loss of crop yield.

L. Agapito et al. (Eds.): ECCV 2014 Workshops, Part IV, LNCS 8928, pp. 105–116, 2015.
DOI: 10.1007/978-3-319-16220-1_8

(a) Sample image from dataset. (b) Field robot used for dataset acquisition.

Fig. 1. Sample image from dataset (a) that was acquired with the autonomous field robot Bonirob (b)

In this paper we consider the use-case of processing top-down looking images of row cultures (organic carrots) with machine vision to capture and extract information that is useful for management and automation of such farming tasks. The image data and annotations made available with this dataset enable the development of solutions for phenotyping problems. Crop / weed discrimination, crop counting, determination of inter-crop spacing or of crop / weed coverage ratios are examples for phenotyping tasks that can be realized and evaluated with this dataset.

From a computer vision perspective the data provided plays an important role: On the one hand the image acquisition process in the agricultural domain is difficult, as it requires complex hardware systems, access to farms and the acquisition must be correctly timed and synchronized to the crop growth cycle (only once a year for many cultures). On the other hand, agricultural experts are needed to define suitable ground truth. That makes this domain different from other problems in computer vision such as object detection in home or street scenes where computer vision researchers can record both data and ground truth more easily themselves. This public dataset allows phenotyping research without the upfront burden of setting up robots, fields and experts.

The dataset comprises field images in top-down view that were acquired with the autonomous field robot Bonirob in an organic carrot farm in 2013 (see Figure 1). All data acquisition was carried out during field tests within the publicly funded project RemoteFarming.1 [3]. The images were captured while the crop was in growth stages where one or more true leaves were present. Some hours after data acquisition the farmer applied manual weed control on this field. Here we consider organic carrots, however similar manual weed control activities are also required for chicory, onions and other cultures. All images are annotated and a ground truth vegetation segmentation mask is available together with crop / weed annotations. Section 3 provides more details about the data, metadata and acquisition conditions.

A concrete example for a phenotyping task which is addressed with this dataset is crop / weed discrimination, for which we provide initial results. A machine vision pipeline is applied and a subset of the images is used together with the ground truth annotations to train a classifier. This classification pipeline is applied to the test images and predicts for each vegetation pixel whether it is part of a crop or weed plant.

To allow comparison of different algorithms we propose evaluation metrics for the vegetation segmentation, plant segmentation and crop / weed discrimination phenotyping tasks.

In summary the contributions of this paper are:

- A dataset of 60 top-down field images of a common culture (organic carrots) with the presence of intra-row and close-to-crop weeds.
- Each image is annotated with a vegetation segmentation mask and crop / weed labels (162 crop plants, 332 weed plants in total).
- The formulation of machine vision and phenotyping problems together with evaluation metrics for future comparison of different approaches.
- Initial results for the crop / weed phenotyping problem of these images.

2 Related Work

In many domains including machine vision, robotics and biology, open datasets are established and play an important role in the scientific community. Public datasets open challenging questions to a wider community and allow direct comparison of different algorithms to the state of the art.

In computer vision there exist many datasets: For example for stereo processing and optical flow, the Middlebury datasets [10,21] and the newer KITTI benchmark [6] are widely used. For image retrieval and object classification larger datasets have been created: For example LabelMe [20], ImageCLEF [16] and the Pascal VOC challenges [5]. In machine learning datasets play an equally important role and a large collection of datasets is available from UCI [2]. Also in robotics, open and public datasets play a major role and for example allow labs without specific robots to do research. KITTI is a dataset for vision based autonomous driving [7], the RGB-D SLAM dataset [25] is a benchmark dataset for simultaneous localization and mapping with depth based vision sensors. Many more datasets exist in all of these domains.

For phenotyping and agricultural tasks however, the availability of datasets is much more limited. In recent years some datasets in the leaf segmentation and classification domain have been published. Söderkvist's Swedish leaf dataset [23] was one of the first available datasets and contains leaf images of Swedish trees. The Flavia dataset by Wu et al. [26] is a newer popular dataset for leaf classification tasks. Kumar et al. developed a Smartphone application for leaf classification called Leafsnap [13] and published their dataset.

The goal of this paper is to provide a real-world field image dataset to the phenotyping, agricultural vision and robotics community. This enables research on perception for data acquisition or treatment in row cultures, such as carrots in early growth stages.

3 Dataset and Problem Description

Figure 2 displays example images from the dataset together with all annotations. The following section describes the content of the dataset, the acquisition parameters as well as the exact format of the image data and metadata.

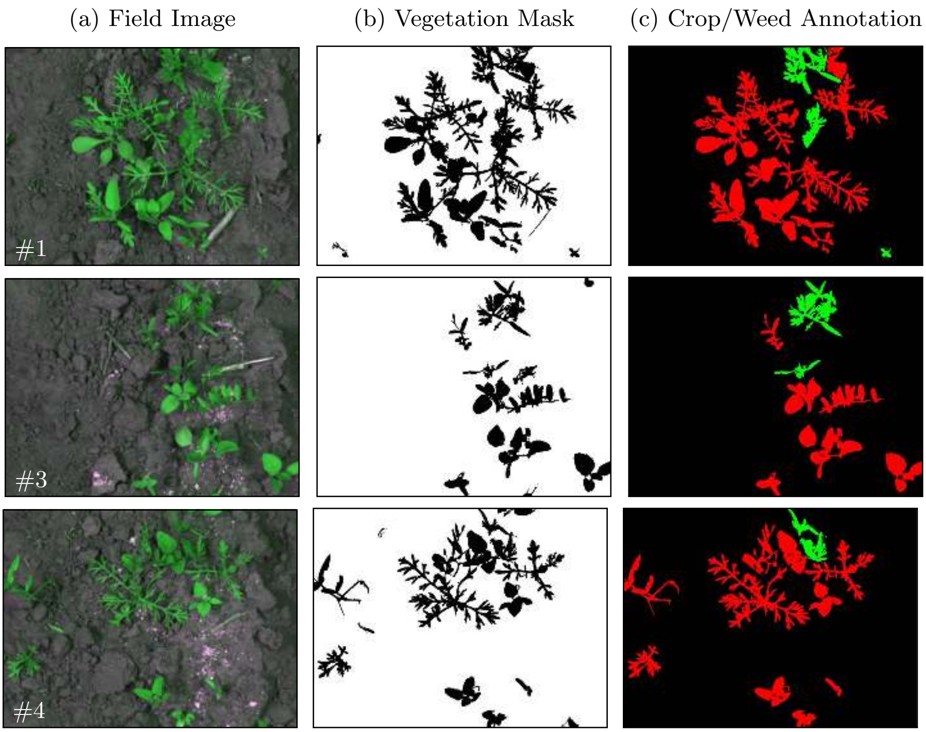

Fig. 2. Sample images from the dataset (a) with ground truth vegetation masks and crop / weed annotations. The annotation images (b) and (c) are supplied for every image of the dataset. Best viewed in color.

3.1 Field Setup and Acquisition Method

The 60 image dataset was captured at a commercial organic carrot farm in Northern Germany in 2013 just before manual weed control was applied. The carrots were grown in single rows on small soil dams. The growth stage of the crop was approximately BBCH 10 – 20 (see [15] for a description of the BBCH plant growth stage scale) and a significant amount of close-to-crop and intra-row weeds was present. Figure 3 describes how the images were selected from five sections in the field. In the agricultural application context where a robot drives along rows, the larger 20 image section at the start of the row is designated as training data, the other sections (40 images) are designated as test set and

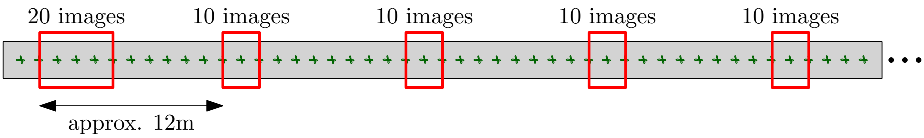

Fig. 3. Schematic overview of a row in the field with annotation of the sections where the dataset images were captured. Near the beginning of the row a section with 20 images was defined; then at a distance of approx. 12 m sections of 10 images each were defined.

Table 1. Extent of the dataset

Parameter	Value
Image count	60
Labeled plant count	494
Labeled crop plant count	162
Labeled weed plant count	332

were spread out across the row to better capture the variability in the field. Subsequent images in the dataset do not overlap and display unique situations to avoid redundant data. Table 1 summarizes the extent of the dataset.

The images were acquired with a camera mounted to the autonomous field robot Bonirob which drove along the carrot row with a speed of ~ 4.5 cm/s. A JAI multi-spectral camera [11] that captures both visible and near-infrared light was used and mounted on the robot. The camera was looking downwards and the area under the robot was shaded and artificially lit to avoid changing lighting conditions. Table 2 describes the camera setup and its configuration. The red (R) and near-infrared (NIR) channels were selected because the spectral characteristics of plants in these channels can be exploited for background removal using vegetation indices [22].

3.2 Dataset and Annotation Format

In addition to the field images the dataset also contains annotations. First, a vegetation mask is provided which masks soil pixels, see Figure 2b. Second, all images were manually annotated by a human expert. The user was asked to mark crop and weed plants / parts with polygons and to assign a type (crop or weed) to each polygon. Note that some areas are not labeled, for example areas with heavy overlap. Figure 2c shows the resulting ground truth crop / weed annotation image when the polygon labels are combined with the vegetation mask. All vegetation pixels that lie inside a polygon inherit the label from the polygon. The label at each pixel is plotted in color code where red denotes weed and green denotes crop. The dataset contains both the polygon information and the crop / weed annotation images as given in Figure 2c. Table 3 summarizes the specific data and file format of the field images and the annotations.

Table 2. Description of camera system and acquisition parameters

Parameter	**Value**
Camera model	JAI AD-130GE [12]
Image resolution	1296 x 966 pixels
Lens	Fujinon TF15-DA-8
Focal length	15 mm
F-number	4
Mean distance to ground (d)	450 mm
Ground resolution	$\sim$ 8.95 pixels/mm
Field of view x (at distance d)	$\sim$ 145 mm
Field of view y (at distance d)	$\sim$ 108 mm

The vegetation masks were derived using the Normalized Differential Vegetation Index (NDVI) [22] that was calculated from the NIR and R image channel. A threshold in NDVI space was selected using Otsu's method [17] given the training images. Then this threshold value was fixed and used to generate the ground truth masks for all images of the dataset.

The crop / weed annotations are given as image (Figure 2c) and in a data format that contains the list of polygons plus one label per polygon (crop / weed). The polygon data is stored in YAML[1] format, see Listing 1. Each YAML file contains a `filename` field and an `annotation` field in which a list of `points` and `type` entries is stored. The `points` field contains the `x` and `y` coordinates of the polygon vertices. The `type` is either crop or weed and defines the plant type.

The crop / weed annotations are also given as polygons because this enables single plant evaluations which are not possible if only an image (Figure 2c) is given. In the annotation image plants of the same type that overlap are no longer separable. Pixels that are covered by more than one polygon with different types are defined as invalid and the plant type is set to unknown.

3.3 How to Get the Dataset

The Crop / Weed Field Image Dataset (CWFID) is available online and can be downloaded from http://github.com/cwfid.

4 Problems and Evaluation Metrics

Field images acquired using a top-down camera system can deliver a lot of information. Nevertheless, their natural setting with different plants growing close together in an unordered scene poses many challenges.

[1] YAML is a data serialization standard which aims to be easy to read for humans. Parsers are available for many programming languages. See yaml.org.

Table 3. Description of dataset and annotation format

Data	**Description**
Field image (Figure 2a)	
Filename	000_image.png
Format	PNG (3 channel), 8 bit
Channels	1 $\mapsto$ Red
	2 $\mapsto$ Near-Infrared
	3 $\mapsto$ Red
Vegetation mask (Figure 2b)	
Filename	000_mask.png
Format	PNG (monochrome), 1 bit
Mapping	Biomass $\mapsto$ 0
	Background $\mapsto$ 1
Crop / weed annotation image (Figure 2c)	
Filename	000_annotation.png
Format	PNG (3 channel), 8 bit
Channels	1 $\rightarrow$ 255 if weed at pixel, 0 otherwise
	2 $\rightarrow$ 255 if crop at pixel, 0 otherwise
	3 $\rightarrow$ always 0
Crop / weed annotation data (Listing 1)	
Filename	000_annotation.yaml
Format	YAML with list of polygon vertices and labels

From a computer vision point of view, these images can be segmented into background and foreground or on a higher level into different objects (for example rows, plants etc.). Furthermore, classification challenges arise including the classification of individual pixels, connected areas or segmented objects. Additionally, many advanced computer vision techniques such as tracking, optical flow etc. can be used to extract information. Some of these tasks overlap with goals of a phenotyping and agricultural image processing point of view.

In the following we are focusing on these more plant specific tasks and formulate four relevant problems:

1. **Vegetation Segmentation**: A binary mask is desired that masks all background soil and residue pixels [14]. Applying this mask results in a vegetation image where only pixels displaying vegetation are non-zero.
2. **Plant Segmentation**: Individual plants should be segmented in the image. This is challenging because plants in the field are growing close together and overlap between plants occurs.
3. **Plant Classification**: Plants or leaves can be classified. Here the use-case of crop / weed discrimination is considered, which results in a two class classification problem. This can be extended to individual species classification.

Listing 1. Definition of the YAML annotations file


```
filename: 000_image.png
annotation:
- type: weed
  points:
    x: [810.0, 841.0, 846.0, 926.0, 956.0, 1054.0]
    y: [225.0, 234.0, 266.0, 338.0, 408.0, 422.0]
- type: crop
  points:
    x: [1070.0, 1055.0, 980.0, 850.0, 844.0]
    y: [626.0, 722.0, 739.0, 658.0, 730.0]
```


4. **Individual Plant Phenotyping**: From the images also information about the phenotype of individual plants can be obtained [19]: This includes the growth stage, plant stem position, biomass amount, leaf count, leaf area and others. Furthermore, crop / weed coverage ratio, inter crop spacing, crop plant count and other derived measurements are of interest to farmers.

For problems 1 - 3, we define evaluation metrics that enable comparison of different approaches when using this dataset. The individual plant phenotyping problems crop plant count and crop / weed coverage ratio can be directly compared to values calculated from ground truth. A definition of metrics for the other phenotyping problems is considered future work and probably requires more annotations.

1. For comparison of different vegetation masks we propose to use the Jaccard index as segmentation accuracy measure (as done in the Pascal VOC challenges [5]), which is defined as intersection over union. This can be expressed in terms of correctly assigned pixels (true positives) and incorrectly assigned pixels (false positives and false negatives):

$$\text{seg. accuracy} = \frac{\text{true pos.}}{\text{true pos.} + \text{false pos.} + \text{false neg.}} \tag{1}$$

   A final score is achieved by averaging the segmentation accuracy over all test images.
2. To evaluate plant segmentation results also the Jaccard index is applied, see Equation (1). The predicted segmentation of a plant (consisting of a set of pixels) is compared with the set of vegetation pixels of the plant in the ground truth annotation. The ground truth vegetation pixels for a single plant are derived by selecting only pixels from the vegetation mask that lie inside the ground truth polygon of the plant. To get a final score the Jaccard index is calculated per plant and then averaged over all plants in the test set.
3. For crop / weed or plant classification, we assume that the classification system returns a full image with per-pixel predictions. Then we propose to

compare the predictions and ground truth pixel-wise and to calculate the following metrics per image: Average accuracy, precision, recall and F1-score [24]. For final results, we propose averaging over the test images. If the prediction also outputs scores and not only binary votes a Receiver Operator (ROC) curve should be plotted.

For tasks that require separate training and test data we propose two splits. First, from an agricultural point of view, we propose a sequential split. Images #1 − 20 located at the beginning of the row are used for training and images #21 − 60 for testing (see Figure 3). This is derived from the real world use-case where system set-up is done at the beginning of the field or row and then performance is expected to be stable during operation.

Second, from a computer vision point of view, we propose a random 66 % train and 33 % test split. Fixed indices for one such split are given in that dataset file `train_test_split.yaml`.

5 Initial Results on Crop / Weed Discrimination

Crop / weed discrimination is an important step towards assessment of crop properties and single plant weed control. Once the type and location of for example crop plants is known, further phenotype measurements can be derived.

Here we provide initial results on the crop / weed discrimination problem on this dataset using the machine vision approach from Haug et al. [8]. In the following, the proposed agricultural test train split is chosen. The 20 training images and vegetation masks are used during the training process, which involves feature extraction using a sliding window approach. For each window position center the corresponding ground truth label is extracted from the ground truth crop / weed annotation. Using the training data (feature vectors) with labels, a Random Forest [4] classifier is trained and applied to the test images of this dataset (images #21 − 60). The predictions of the Random Forest classifier are post-processed and the output of the plant classification system is a predicted crop / weed image similar to the ground truth image.

Figure 4 displays the crop / weed predictions next to a ground truth image from the dataset. In both the ground truth image and the predicted image each vegetation pixel is plotted in color code, where red denotes weed and green denotes crop. A border of 40 pixels is masked and was ignored during evaluation, as this approach does not predict the plant type at the edges of the image.

To quantitatively analyze the performance of this approach to crop / weed discrimination, pixel-wise comparison of ground truth image and prediction is applied. Table 4 summarizes the proposed per-pixel metrics averaged over all test images.

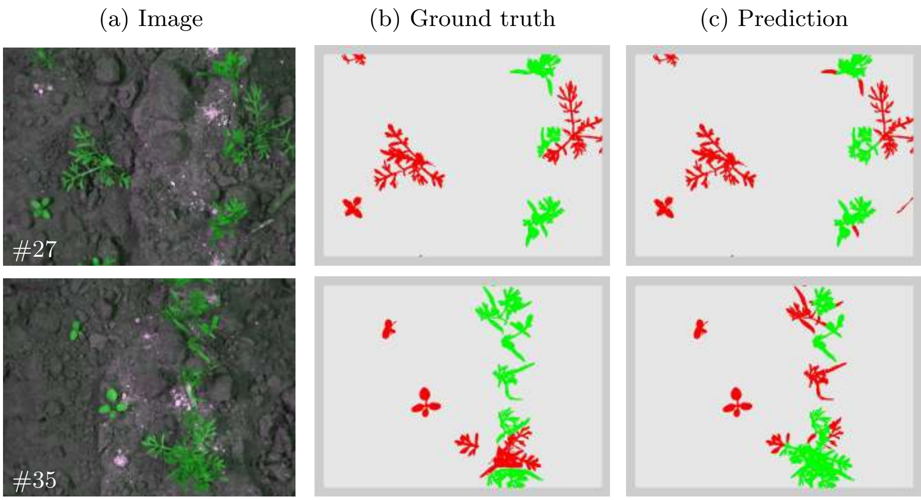

Fig. 4. Image, ground truth and crop / weed prediction for two test images. Red color denotes weed and green color denotes crop. Best viewed in color.

Table 4. Results of crop / weed classification when comparing per-pixel predictions of test images with the ground truth

Metric	**Result**
Average Accuracy	85.9 %
Precision	79.6 %
Recall	80.8 %
F1-score	80.2 %

6 Conclusions

This paper proposes a crop / weed field image dataset for phenotyping and machine vision problems in agriculture. Field images of carrots were acquired on a commercial organic farm in early crop growth stage, where close-to-crop and intra-row weeds were present. Such images pose both phenotyping and machine vision related questions that – if solved – allow the automation of manual and cost intense tasks including for example weed control.

The data is fully annotated by experts. Initial results on crop / weed discrimination report an average accuracy and F1-score of 85.9 % and 80.2 % respectively. This indicates that automation of such tasks is probably feasible, however difficult and needs more research.

Finally, we propose evaluation metrics for segmentation and classification tasks to encourage other groups to use this dataset and compare results. We hope that this increases progress in this domain where data acquisition requires extensive setups, experts with agricultural knowledge are needed to generate ground truth and availability of public datasets is very limited.

In the future, this dataset can be enlarged with more images from another field or growth season and additional ground truth can be defined for the individual plant phenotyping problems.

Acknowledgments. The authors thank the following colleagues for their comments and help with the acquisition of the dataset: Wolfram Strothmann, Fabian Sellmann, Arno Ruckelshausen, Susanne Fittje, Frederik Langsenkamp, Maik Kohlbrecher, Dieter Trautz (University of Applied Sciences Osnabrück), Waldemar Bangert, Florian Rahe (Amazone), Andreas Michaels, Slawomir Grzonka, Peter Biber, Hannes Becker, Amos Albert (Bosch).

The project RemoteFarming.1 is partially funded by the German Federal Ministry of Food, Agriculture and Consumer Protection (BMELV).

References

1. Åstrand, B., Baerveldt, A.J.: An agricultural mobile robot with vision-based perception for mechanical weed control. Autonomous Robots **13**(1), 21–35 (2002)
2. Bache, K., Lichman, M.: UCI machine learning repository (2013). http://archive.ics.uci.edu/ml
3. Bangert, W., Kielhorn, A., Rahe, F., Albert, A., Biber, P., Grzonka, S., Haug, S., Michaels, A., Mentrup, D., Hänsel, M., Kinski, D., Möller, K., Ruckelshausen, A., Scholz, C., Sellmann, F., et al.: Field-robot-based agriculture: RemoteFarming. 1 and BoniRob-Apps. VDI. Agricultural Engineering **2013**, 439–446 (2013)
4. Breiman, L.: Random forests. Machine Learning **45**(1), 5–32 (2001)
5. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International Journal of Computer Vision **88**(2), 303–338 (2010)
6. Geiger, A., Lenz, P., Stiller, C., Urtasun, R.: Vision meets robotics: The kitti dataset. The International Journal of Robotics Research **32**(11), 1231–1237 (2013)
7. Geiger, A., Lenz, P., Urtasun, R.: Are we ready for autonomous driving? the kitti vision benchmark suite. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3354–3361. IEEE (2012)
8. Haug, S., Michaels, A., Biber, P., Ostermann, J.: Plant classification system for crop/weed discrimination without segmentation. In: 2014 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1142–1149. IEEE (2014)
9. Hemming, J., Rath, T.: Computer-vision-based weed identification under field conditions using controlled lighting. Journal of Agricultural Engineering Research **78**(3), 233–243 (2001)
10. Hirschmuller, H., Scharstein, D.: Evaluation of cost functions for stereo matching. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–8. IEEE (2007)
11. JAI: Datasheet: JAI AD-130GE 2 CCD Multi-spectral Camera., Document version: March 2012

12. JAI: User's Manual: JAI AD-130GE 2CCD Multi-Spectral Camera., Document version: 1.1 (2012)
13. Kumar, Neeraj, Belhumeur, Peter N., Biswas, Arijit, Jacobs, David W., Kress, WJohn, Lopez, Ida C., Soares, João V.B.: Leafsnap: A Computer Vision System for Automatic Plant Species Identification. In: Fitzgibbon, Andrew, Lazebnik, Svetlana, Perona, Pietro, Sato, Yoichi, Schmid, Cordelia (eds.) ECCV 2012, Part II. LNCS, vol. 7573, pp. 502–516. Springer, Heidelberg (2012)
14. McCarthy, C., Hancock, N., Raine, S.R.: Applied machine vision of plants: a review with implications for field deployment in automated farming operations. Intelligent Service Robotics **3**(4), 209–217 (2010)
15. Meier, U.: Growth stages of mono-and dicotyledonous plants. BBCH monograph. German Federal Biological Research Centre for Agriculture and Forestry, Berlin (2001)
16. Müller, H., Clough, P., Deselaers, T., Caputo, B., CLEF, I.: Experimental evaluation in visual information retrieval. The Information Retrieval Series 32 (2010)
17. Otsu, N.: A threshold selection method from gray-level histograms. Automatica **11**(285–296), 23–27 (1975)
18. Ruckelshausen, A., Biber, P., Dorna, M., Gremmes, H., Klose, R., Linz, A., Rahe, F., Resch, R., Thiel, M., Trautz, D., et al.: Bonirob - an autonomous field robot platform for individual plant phenotyping. Precision Agriculture **9**, 841 (2009)
19. Ruckelshausen, A., Busemeyer, L., Klose, R., Linz, A., Moeller, K., Thiel, M., Alheit, K., Rahe, F., Trautz, D., Weiss, U.: Sensor and system technology for individual plant crop scouting. In: International Conference on Precision Agriculture (ICPA) (2010)
20. Russell, B.C., Torralba, A., Murphy, K.P., Freeman, W.T.: Labelme: a database and web-based tool for image annotation. International Journal of Computer Vision **77**(1–3), 157–173 (2008)
21. Scharstein, D., Szeliski, R.: A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. International Journal of Computer Vision **47**(1–3), 7–42 (2002)
22. Scotford, I., Miller, P.: Applications of spectral reflectance techniques in northern european cereal production: a review. Biosystems Engineering **90**(3), 235–250 (2005)
23. Söderkvist, O.: Computer Vision Classification of Leaves from Swedish Trees. Master's thesis, Linköping University, Sweden (2001)
24. Sokolova, M., Lapalme, G.: A systematic analysis of performance measures for classification tasks. Information Processing & Management **45**(4), 427–437 (2009)
25. Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D.: A benchmark for the evaluation of rgb-d slam systems. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 573–580. IEEE (2012)
26. Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.X., Chang, Y.F., Xiang, Q.L.: A leaf recognition algorithm for plant classification using probabilistic neural network. In: 2007 IEEE International Symposium on Signal Processing and Information Technology, pp. 11–16. IEEE (2007)