The P2S2 segmentation dataset: annotated in-field multi-crop RGB images acquired under various conditions

Simon Madec[1], Kamran Ifran[1], Etienne David[1,2], Kaaviya Velumani[1,3], Gaetan Daubige[2], Jeremy Labrosse[3] , Wei Guo[4],Marie Weiss[1], Frederic Baret[1]

[1]INRA, UMT-CAPTE, EMMAH UMR1114, Avignon, France
[2]Arvalis, Avignon, France
[3]Hiphen SAS, Avignon, France
[4]The University of Tokyo · Japan

Abstract

Images play a vital role in crop phenotyping. Pixel-wise classification (into vegetation/background) or semantic segmentation is a critical step in the computation of several canopy state variables. Current state of the art methodologies based on convolutional neural networks are trained on data acquired under controlled environments. These models are unable to generalize to real-world dataset and hence need to be fine-tuned using new labels. This motivated us to create the P2S2 segmentation dataset – a collection of multi-crop RGB images from different acquisition conditions. We present here the dataset and state of the art results.

Keywords: Dataset, Segmentation, Convolutional Neural Networks, RGB, Crop

1 Introduction

Deep learning and convolutional neural networks (CNNs) have recently demonstrated their huge potential in plant image segmentation [1]. However, the training and validation images are often acquired in laboratory and under controlled illumination conditions [2]–[3]. The evaluation of the method performances is also often limited to specific conditions including species or cultivars, crop stage, and illumination conditions. To overcome this issue, we propose the P2S2 dataset of annotated images: it was initially acquired for the validation of vegetation products derived from the Sentinel 2 satellite. It is composed of 75 images acquired over nine crops and different stages/conditions. We first describe our annotation strategy to build an accurate dataset. We then use the P2S2 segmentation dataset to evaluate the performances of a CNN trained with already existing datasets from the literature. We hope that this publicly available dataset will help crop phenotyping research to overcome the bottleneck in robust segmentation model building.

2 Materials and methods

2.1 Image collection: The data collection was carried out in four sites in France and Belgium chosen for their varied climatic and soil conditions. These four sites were cultivated with nine crop species - wheat, rapeseed, maize, sunflower, sugar beet, rice, potato, soybean and grassland. We considered five acquisition dates across the growing season in such a way to cover as many growth stages as possible. Downward looking digital RGB images were acquired at the ground level. Approximately 2400 images of size 6000 x 4000 pixels were collected through this experiment with a spatial resolution of 0.2mm. From these images, we picked up 20 random patches of 512 x 512 pixels for each crop species. A maximum of four patches before and after the appearance of organs was selected. This led to a total of 75 images

2.2 Annotation strategy: Thirteen experts precisely annotated the 75 images using a custom improved version of the JavaScript annotation tool provided by [4]. To avoid annotation bias due to subjectivity, at least two independent experts reviewed each annotated image. The dataset was annotated following a simple rule: labeling all the pixels belonging to a plant as vegetation (including flowers, spikes, and dried leaves) and the rest as background.

Figure 1 Examples of images and their corresponding annotated masks

2.3 CNN architecture:
We built a CNN inspired from the U-Net convolution-deconvolution architecture. We used ResNet with 18 layers, initialized on the ImageNet dataset, as the backbone architecture. The training dataset was composed of five ready-to-use datasets that included the images and their corresponding vegetation/background mask. These datasets correspond to a range of vegetation: CVPP dataset (top view of rosette plants), Easy-PCC (rice and wheat fields), wheat, carrot and weeds, and one other dataset of wheat (unpublished) [5]–[7]. This constituted 1400 images, from which 15% was withheld for validation.

2.4 Evaluation metrics:
The CNN model was tested on the P2S2 segmentation dataset and its performances were also compared with a random forest (RF) model. The RF classifier was trained with the RGB pixel values, and we used 100 trees in the final estimator parameters. The fraction of vegetation and the mean IOU from these two approaches were compared.

3 Results and Discussion

3.1 CNNs achieve comparable performance to RF:

Both approaches achieved a relative error of 20% on the P2S2 dataset. The lowest performances were achieved on crop species that were not selected in the training database. For the CNN, the performances are affected by the spatial resolution and the sharpness of the images. Regarding the RF model, the main limitations are due to strong illumination conditions, soil appearance and the presence of non-green vegetation.

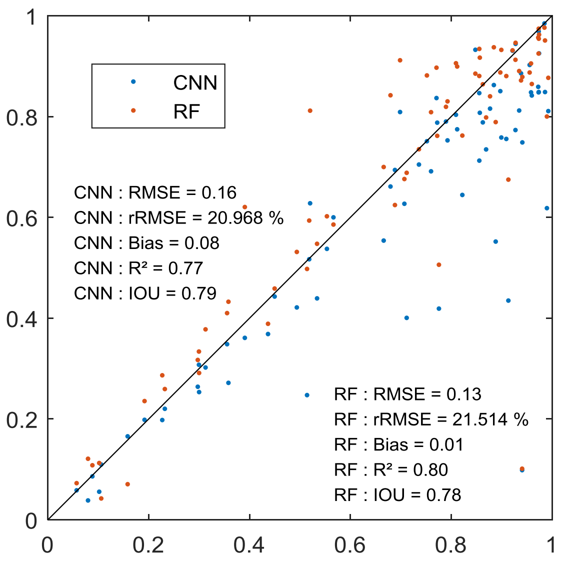

Figure 2: Vegetation cover estimates by CNN and RF

4 Conclusion and Perspective

We introduced a new, carefully annotated, precise, diverse image dataset acquired in a realistic environment. This challenging dataset is designed to improve the segmentation of vegetation images by building robust models. Moreover, it could be used as a common benchmark for future studies on vegetation segmentation It was used to evaluate a CNN approach trained on existing datasets from the literature. Results showed quite poor performances of this method due to the discrepancies between the training and the P2S2 dataset.This highlights the overall-value of this new dataset. Future work will focus on data augmentation and domain adaptation.

Acknowledgement

We acknowledge the financial support of the Centre National d’Etudes Spatiales (CNES) through a research grant under the program TOSCA (Terre solide, Océan, Surfaces Continentales, Atmosphère) for the acquisition of in situ measurements. We are also very grateful to the participants of the in situ measurements over the different sites.

References

[1] H. Scharr *et al.*, “Leaf segmentation in plant phenotyping: a collation study,” *Mach. Vis. Appl.*, vol. 27, no. 4, pp. 585–606, 2016.
[2] P. Lottes, J. Behley, A. Milioto, and C. Stachniss, “Fully Convolutional Networks with Sequential Information for Robust Crop and Weed Detection in Precision Farming,” *ArXiv Prepr. ArXiv180603412*, 2018.
[3] J. R. Ubbens and I. Stavness, “Deep Plant Phenomics: A Deep Learning Platform for Complex Plant Phenotyping Tasks,” *Front. Plant Sci.*, vol. 8, 2017.
[4] P. Tangseng, Z. Wu, and K. Yamaguchi, “Looking at outfit to parse clothing,” *ArXiv Prepr. ArXiv170301386*, 2017.
[5] H. Scharr, M. Minervini, A. Fischbach, and S. A. Tsaftaris, “Annotated image datasets of rosette plants,” 2014.
[6] W. Guo, B. Zheng, T. Duan, T. Fukatsu, S. Chapman, and S. Ninomiya, “EasyPCC: Benchmark Datasets and Tools for High-Throughput Measurement of the Plant Canopy Coverage Ratio under Field Conditions,” *Sensors*, vol. 17, no. 4, Apr. 2017.
[7] S. Haug and J. Ostermann, “A crop/weed field image dataset for the evaluation of computer vision based precision agriculture tasks,” in *European Conference on Computer Vision*, 2014, pp. 105–116.