59. Plant phenotyping on-demand: an integrative web-based framework using drones and participatory sensing in greenhouses

Precision agriculture '21

Frangulea, M.; Pantos, C.; Giuffrida, V.; Valente, J.

https://doi.org/10.3920/978-90-8686-916-9_59

59. Plant phenotyping on-demand: an integrative web-based framework using drones and participatory sensing in greenhouses

M. Frangulea[1], C. Pantos[1], V. Giuffrida[2], and J. Valente[1*]
[1]*Information Technology group, Wageningen University, the Netherlands; joao.valente@wur.nl*
[2]*School of Computing, Edinburgh Napier University, United Kingdom*

Abstract

A tool for plant phenotyping is proposed to aid users in analyzing data on-demand. This tool is web-based and runs deep learning models. The current study focuses on the development of this tool, as well as obtaining a plant dataset to train a neural network. Furthermore, smartphone and drone imagery are used to test the derived model. The results demonstrate how data generalization can be reached through participatory sensing. Finally, drones show potential as being a fast solution for acquiring sensory information within greenhouses.

Keywords: kale, smartphone, drone, greenhouse, artificial intelligence

Introduction

Plant phenotyping represents the observable characteristics of the physiological and biochemical properties of a plant, as well as its anatomical and ontological traits (Guo & Zhu, 2015). Plant phenotyping can be used to improve the decision-making process in agriculture. Furthermore, breeders benefit from plant phenotyping as they can enhance the genotype selection. Since plant phenotyping is useful in understanding the functioning of crops, it also helps in developing and calibrating crop models. Traditionally, characterizing the plant phenotype required a significant number of experienced individuals recording plant characteristics manually: a labour intensive and often invasive process. More recently, plant phenotyping evolved into a digital, non-invasive discipline (Costa *et al.*, 2019). With this evolution, plant phenotyping is becoming useful to plant breeding and selection. Multiple traits such as yield and stress adaptation are now measured in order to optimize the required parameters for environmental adaptation (Fiorani & Schurr, 2013). To perfect plant phenotyping, multiple study domains need to be combined. From biology to computer science, today's challenge is to merge these disciplines under one common language (Costa *et al.*, 2019). The aim of the current study is to create a bridge between plant and information sciences by facilitating the data analysis process. This can be achieved through an online system based on deep learning algorithms. Ultimately, plant scientists and horticulturists will benefit from this system as they will not be required to be proficient in another discipline. Data collected will be valuable regardless of researchers' computer abilities. In plant sciences, applications of deep learning include classifying plants and deriving relevant phenotypes that may explain more about the plant physiology and status (Christin *et al.*, 2019).
There are many image analysis tools designed for plant phenotyping, most of which are available as downloadable software. Whether they are standalone executables or plugins, installing the tools is required. Out of 179 tools covered in a plant phenotyping database, only one of them is web based (Lobet *et al.*, 2013). Furthermore, that tool only addresses a specific niche: counting stomata (Fetter *et al.*, 2019). The web-based application proposed in this paper represents a starting point to bridging the gap between farmers and plant phenotyping technology.
The aim of the study is to bring current technology closer to the farmer following several research objectives. The first one is designing an integrative framework to aid farmers into collecting and

DOI 10.3920/978-90-8686-916-9_59, © Wageningen Academic Publishers 2021

analyzing data. The second objective is gathering an image dataset using two cultivars of a rosette plant. The image dataset is to be released as an open-source package for further research. The third research objective is using the acquired dataset to train a deep neural network for plant phenotyping, deriving an improved model. Furthermore, the fourth objective is to test smartphone and drone imagery on the resulting model using the web application developed within the study. Finally, the feasibility of using drones for data acquisition inside greenhouses is evaluated.

Materials and methods

Participatory sensing approach

Participatory sensing is an approach through which individuals contribute to a body of knowledge by using a mobile device (Tilak, 2013) to upload information. This approach helps both the user and the expert. For example, farmers can obtain phenotype information by collecting data in the greenhouse or the field and uploading it to a server for processing. The information can be further used to improve plant phenotyping models. A scheme of participatory sensing is illustrated in Figure 1.

Experimental setup and data acquisition

Two kale varieties were grown inside the greenhouse of Wageningen University & Research, in two stages. Each stage contained half of the plants and were grown for 28 days: June 16th – July 14th (Stage 1) and July 10th – August 7th (Stage 2). In total, 160 plants were grown per variety, totaling 320 plants. Eight days after seeding, the plants reached the necessary growth stage to begin the dataset acquisition phase. Thus, all plants were photographed individually for 20 consecutive days, resulting in a 6,400-image dataset. When taking the images, a high degree of standardization was intended. Thus, the phone holder was used to ensure the same angle and height. Furthermore, the pots were placed in the same position without being rotated between shots (Figure 2).

Sensory information was obtained by taking images of plants. A standardized dataset was obtained with a OnePlus 6T (OnePlus Technology, China, released 2018) mobile phone. In order to account for diversity in mobile phones owned by farmers as well as different acquisition platforms and conditions, images from different angles and of different quality were taken using two mobile phones – iPhone 6S (Apple Inc., USA, released 2015), Samsung SGH-T679 (Samsung Group, South Korea, released

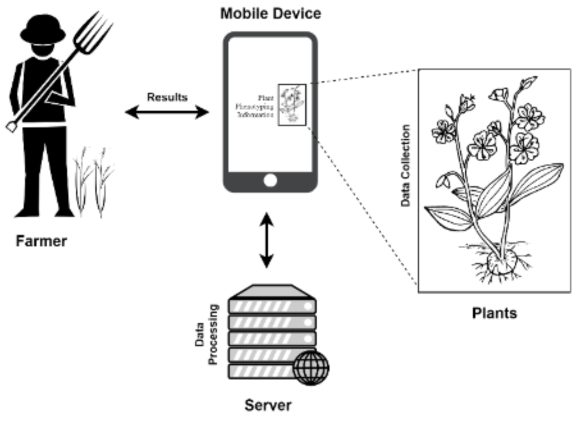

Figure 1. Participatory sensing scheme. Data are collected using a mobile device or drone and uploaded to the server where it is processed. Results are displayed on the device.

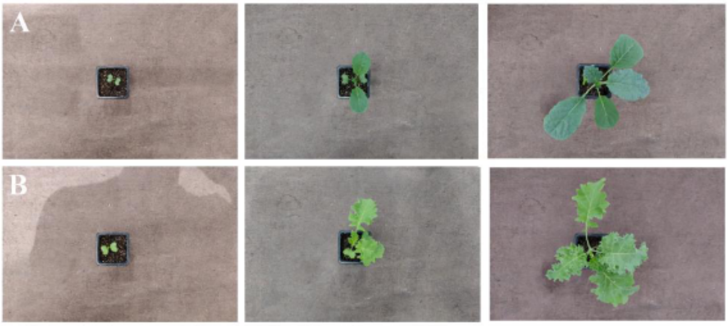

Figure 2. Examples of dataset images. From top to bottom: day 1, day 10, day 20 of (a) *Brassica oleracea* var. *acephala* 'Nero di Toscana' and (b) *B. oleracea* var. *sabellica.*

2011) and a Parrot Anafi (Parrot SA, France, released 2018) drone (Figure 3). The mobile phones were chosen to compare performance between different smartphone generations.

Image classification

The dataset was used for training a deep neural network for leaf counting, one of the more important traits of plant phenotyping which gained a lot of attention recently (Scharr *et al.*, 2016). All images were rescaled due to computer processing constraints. Thus, the open source Pheno-Deep Counter architecture (only the RGB branch) was chosen for its high performance (Giuffrida *et al.*, 2018). This architecture uses the ResNet50 convolutional neural network as feature extractor, with output fed to three fully connected layers to obtain the final leaf counting prediction. The ResNet50 backbone was initialized with ImageNet weights and then fine-tuned with the standardized images of kale. Specifically, the first half of the images was used for training, while the second half for validation and testing. The resulting model was used for evaluating the mobile phones and drone imagery in leaf counting. For the smartphones, fifty images were randomly chosen from a subsequent dataset for each variety. The drone was operated at two heights, the lower one at approximately one meter and the higher height at two meters. Furthermore, drone imagery was cropped to fit one pot per image, ten images per height being analysed. To assess the dataset, evaluation metrics were used as in Giuffrida *et al.* (2018): difference in count [dic]: average of the difference between the ground

Figure 3. Drones usage for data collection.

truth and the algorithmic prediction; absolute difference in count [|dic|]: similar to dic, but each difference is computed in absolute value; Mean Square Error [MSE]: average of the squared difference between the ground truth and the prediction; Percentage Agreement [PA]: number of times (in percentage) the model correctly predicts the leaf count. Furthermore, data significance was assessed using Student's independent t-test at a significance level [*P*] of 0.05. The [t] statistic is presented, as well as the mean [M] and standard deviation [SD] of each tested group.

Website application

A website application was built as part of the integrative framework (available at https://saidlab.wur.nl/phenotyping). The application was developed using Django within its architecture, a high-level Python web framework (Anonymous, 2021). To take advantage of the machine learning deployment capabilities of Django, the web application was hosted on Amazon Web Services. A registration system was set up on the website. User registration is free but not required to use the tool. Thus, registered users benefit from uploading and processing multiple images at once (a maximum of ten). The framework was designed to support various plant phenotyping models. Currently, the web application runs the model trained within the current project. Once the images are processed by the server, users are shown the results on the web page. Registered users can also access the analysis history on a separate page (Figure 4).

Results

Neural network training

The network was trained in two different conditions: intra- and inter-variety. In the intra-variety, the network was trained and tested on the same kale species. Specifically, it was trained with the var. *acephala* and evaluated with images of the same variety. Likewise, the model was trained and tested with the var. *sabellica*. As shown in Table 1, training with either of the varieties of kale gave encouraging results, showing a MSE of approx. 1. Furthermore, following the same training paradigm as in Dobrescu *et al.* (2017), the neural network was also trained by combining both varieties. As shown in Table 1, the MSE dropped to 0.82, resulting in a more accurate model. This is in line with

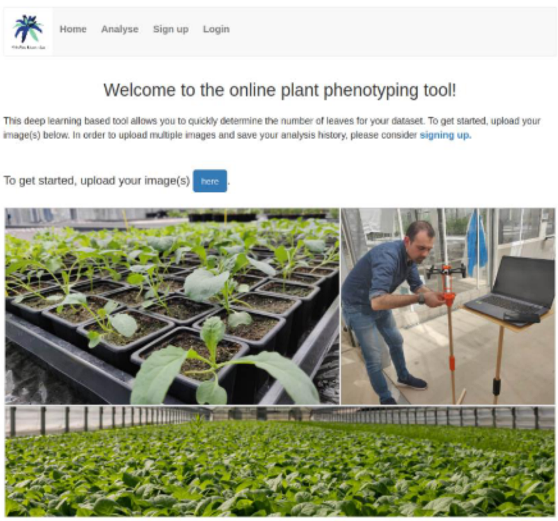

Figure 4. Web application landing page.

Table 1. Leaf counting testing intra- and inter-species results for *B. oleracea* var. *acephala* 'Nero di Toscana' and *B. oleracea* var. *sabellica*. dic and |dic| reports in parenthesis the standard deviation.

	Trained on	Tested on	dic	|dic|	MSE	PA
Intra variaety	var. *acephala*	var. *acephala*	-0.2 0 (1.0)	0.75 (0.71)	1.1	40
	var. *sabellica*	var. *sabellica*	-0.41 (0.92)	0.76 (0.66)	1.0	35
	All	All	-0.03 (0.42)	0.18 (0.38)	0.82	82
Inter variaety	var. *acephala*	var. *sabellica*	-0.21 (1.1)	0.82 (0.72)	1.2	62
	var. *sabellica*	var. *acephala*	0.09 (0.87)	0.58 (0.65)	0.76	50

the findings shown in Dobrescu *et al.* (2017) and Giuffrida *et al.* (2018). Although the two varieties of kale exhibit difference in visual appearance, the model takes advantage of the fact that it is learning from a bigger set of images, making it more accurate. As such, PA goes to 82% (compared to 35% and 40% for the two previous experiments respectively).

In the inter-variety experiment, the goal was to evaluate the generalizability of the model with the kale dataset. Specifically, the trained models from the intra-species experiment (when trained with one variety only) was tested with the other variety. For instance, the neural network was tested with the var. *sabellica* using the model trained on var. *acephala* – and vice versa. The results of these experiments are also reported in Table 1. As expected, the model trained on var. *achephala* underperforms when tested against var. *sabellica* images. However, the PA has almost a 2-fold improvement. This means that the network predicts the correct number of leaves in 62% of cases. Consequently, when the network makes a prediction error, such an error is (on average) higher (this justifies an MSE 1.2 compared to 1.0). When the network was tested against the var. *acephala*, using the model trained on var. *sabellica*, the prediction error dropped significantly, compared to the model both trained (and tested) on var. *acephala* – and the PA increased to 62%.

Other mobile phones and drone images

Smartphone imagery was tested using the model trained on both varieties. The Samsung SGH-T679 scored lower MSE for both varieties, PA being higher for the *B. oleracea* var. *acephala* compared to the iPhone 6S (Table 2).

The iPhone 6S performed poorly because of the differences in image quality. Due to the increased number of outlying predictions, the iPhone scored a high MSE and low PA for both varieties. *B. oleracea* var. *acephala* showed significant differences ($t(49)=-4.95$, $P<0.00001$) between the observed ($M=4.74$, $SD=0.36$) and the predicted values ($M=5.54$, $SD=0.95$). *B. oleracea* var. *sabellica* also recorded a significant difference ($t(49)=-2.21$, $P=0.15$) comparing the observed ($M=5.06$, $SD=0.67$) and predicted results ($M=5.42$, $SD=0.66$).

Table 2. Effect of smartphone model on leaf counting for *B. oleracea* var. *acephala* 'Nero di Toscana' and *B. oleracea* var. *sabellica*.

Model	var. *acephala*		var. *sabellica*	
	MSE	PA	MSE	PA
iPhone 6S	1.48	0.38	1.4	0.42
Samsung SGH-T679	0.46	0.72	0.44	0.31

The Samsung SGH-T679 showed a high degree of PA. There were no significant differences between the values for either variety. Thus, no significant difference ($t(49)=-0.87$, $P=0.19$) was observed for *B. oleracea* var. *acephala* when comparing the observed values (M=4.9, SD=0.17) to those predicted (M=5, SD=0.49). Furthermore, no significant differences ($t(49)=-0.64$, $P=0.26$) were recorded for *B. oleracea* var. *sabellica* between the observed (M=5.08, SD=0.61) and predicted (M=5.16, SD=0.18) values.
B. oleracea var. *acephala* 'Nero di Toscana' was tested using drone imagery, on the model trained on both varieties. While PA was low, the MSE for both heights was below one. This demonstrates that picture quality was satisfactory, the Parrot Anafi outperforming the iPhone 6S when comparing MSE. Both MSE and PA scored poorer for the lower height (Table 3).
In all incorrectly predicted instances, the error margin was of one unit. This was probably due to the plant changes that are introduced by the drone rotor wind speed. It was observed that lower flying heights can interfere with the positioning of the leaves, ultimately corrupting the dataset. In the current study, to avoid plant changes, images were taken from the side instead of the top.

Discussion

In contrast to that reported in the literature (e.g. Giuffrida *et al.*, 2018), testing the models with the different variety of kale resulted in either a mild loss or a significant improvement of performance. This phenomenon is mainly related to the quality of the images. Specifically, the hand-held device used to acquire the standardized dataset took high-resolution images (approximately 4,600×3,400 pixels). However, the model cannot handle such big images due to memory constraints of the graphics processing unit and, thus, it needs to be rescaled to 320×320 (Giuffrida *et al.*, 2018). This reduction of resolution is enough to guarantee robust predictions (as demonstrated in Table 1), but dramatically reduces the details in the plants, making the two varieties look similar at that reduced scale. For this reason, the inter-species experiment resulted in consistent results. This evidence is partially in contrast to what is shown in Giuffrida *et al.* (2018), although, in that case, the authors dealt with the inter-species (e.g. different plant species), rather than inter-variety. In conclusion, it can be said that the model presented gives robust results on var. *sabellica* when trained on the var. *acephala* (and vice versa). However, to obtain the best performance, the network needs to be trained with more data. As such, as demonstrated in Table 1, the network predicts the exact leaf count in more than 80% of cases when trained on both varieties, which is in agreement with what is shown in Dobrescu *et al.* (2017).
A need for open-source datasets is often reported in the literature (Tsaftaris & Scharr, 2019). Furthermore, data generalization is one of the current goals of the scientific community (Giuffrida *et al.*, 2018). With the different devices used for testing plant imagery, this study advances the research towards a general model. While not consistent, the mobile phones showed promising results. The Samsung SGH-T69, a ten-year-old model, outperformed the newer phone on the market. This represents an advantage for the farmer as older technology is not a barrier between them and online plant phenotyping. Furthermore, a high MSE and low PA were noted for the iPhone 6S. The differences within intrinsic camera settings, as well as hardware variations could be the reasons why this smartphone was outperformed by the Samsung (Cobârzan *et al.*, 2015). Image processing

Table 3. Parrot Anafi models for *B. oleracea* var. *acephala* 'Nero di Toscana'.

Model	MSE	PA
Higher height	0.4	0.6
Lower height	0.9	0.1

might have also altered the results, color saturation being shown to influence predictions (Castro *et al.*, 2019). Previous studies showed the importance of minimizing differences between smartphone devices for improved results (Pichon *et al.*, 2020). Thus, using the leaf counting model, future studies should explore participatory sensing by testing a larger number of smartphones. This will improve the understanding of the differences between devices and how distinct smartphone generations influence results. Aside from the turbulence issue influencing the positioning of the leaves, drone imagery is a potential solution for the future of plant phenotyping. While having a low PA, the MSE was below one for both varieties and heights. All predictions indicated a maximum inaccuracy of one leaf. Collecting data using drones inside greenhouses can therefore be a fast and reliable option. Further research is required to obtain a general model for leaf counting. The kale dataset is available at https://doi.org/10.5281/zenodo.4315437 under the Creative Commons Attribution 4.0 International license.

Conclusions

The current study proposed an online tool for plant phenotyping for non-expert users. This was achieved through an integrative framework which has a web application at its center. Through this application, farmers can process the data gathered through participatory sensing. Different plant phenotyping models can be used within the framework, the tool currently running a leaf counting model. A plant dataset was obtained for two purposes: training the network based on the Pheno-Deep Counter architecture and releasing the dataset as a standalone resource for future research. A leaf counting model was tested using mobile phone and drone imagery. The results showed potential for data generalization. Finally, drone imagery could be used in tomorrow's plant phenotyping, drones potentially being a fast solution for data collection inside greenhouses.

Acknowledgements

The authors would like to thank Wageningen Unifarm and the greenhouse manager Andre Maassen for supporting the experiments.

References

Anonymous (2021). Django – why Django? https://www.djangoproject.com/start/overview.

Castro, W., Oblitas, J., De-La-Torre, M., Cotrina, C., Bazán, K., and Avila-George, H. (2019). Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces. IEEE Access, 7, 27389-27400.

Cobârzan, C., Hudelist, M., Schoeffmann, K., and Primus, M. (2015). Mobile image analysis: android vs iOS. In International Conference on Multimedia Modeling (pp. 99-110). Cham, Switzerland: Springer.

Costa, J.M., Marques da Silva J., Pinheiro, C., Barón, M., Mylona, P., Centritto, M., *et al.* (2019). Opportunities and Limitations of Crop Phenotyping in Southern European Countries. Frontiers in Plant Science. 10, 1125.

Christin, S., Hervet, E., and Lecomte, N. (2019). Applications for deep learning in ecology. Methods in Ecology and Evolution, 10(10), 1632-1644.

Dobrescu, A., Giuffrida, V., and Tsaftaris, S. (2017). Leveraging multiple datasets for deep leaf counting. In Proceedings of the IEEE International Conference on Computer Vision Workshops, pp. 2072-2079.

Fetter, K., Eberhardt, S., Barclay, R., Wing, S, and Keller, S. (2019). StomataCounter: a neural network for automatic stomata identification and counting. New Phytologist, 223(3), 1671-1681.

Fiorani, F., and Schurr, U. (2013). Future scenarios for plant phenotyping. Annual Review of Plant Biology, 64, 267-291.

Giuffrida, M. V., Doerner, P., and Tsaftaris, S. (2018). Pheno-deep counter: a unified and versatile deep learning architecture for leaf counting. The Plant Journal, 96(4), 880-890.

Guo, Q., and Zhu, Z. (2006). Phenotyping of plants. In Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation, Hoboken, NY, USA: John Wiley. pp 1-15.

Lobet, G., Draye, X., and Périlleux, C. (2013). An online database for plant image analysis software tools. Plant Methods 9, 38.

Pichon, L., Taylor, J., and Tisseyre, B. (2020). Using smartphone leaf area index data acquired in a collaborative context within vineyards in southern France. OENO One, 54(1), 123-130.

Scharr, H., Minervini, M., French, A., Klukas, C., Kramer, D., Liu, X., *et al.* (2016). Leaf segmentation in plant phenotyping: a collation study. Machine vision and applications, 27(4), 585-606.

Tilak, S. (2013). Real-world deployments of participatory sensing applications: Current trends and future directions. International Scholarly Research Notices, 2013, 583165.

Tsaftaris, S., and Scharr, H. (2019). Sharing the right data right: a symbiosis with machine learning. Trends in Plant Science, 24(2), 99-102.