Smart Agricultural Technology 7 (2024) 100383

Contents lists available at ScienceDirect

Smart Agricultural Technology

journal homepage: www.journals.elsevier.com/smart-agricultural-technology

Automatic trait estimation in floriculture using computer vision and deep learning

Manya Afonso [a,*], Maria-João Paulo [a], Hubert Fonteijn [a], Mary van den Helder [b], Henk Zwinkels [b], Marcel Rijsbergen [c], Gerard van Hameren [c], Raoul Haegens [c], Ron Wehrens [a]

[a] Wageningen University and Research, Wageningen, the Netherlands
[b] Floricode, Roelofarendsveen, the Netherlands
[c] Naktuinbouw, Roelofarendsveen, the Netherlands

ARTICLE INFO

Editor: Spyros Fountas

Keywords:
Floriculture
Ornamentals
Variety registration
Computer vision
Deep learning

ABSTRACT

Registration of new varieties of ornamental flowers is an important process in protecting plant breeders' intellectual property as well as consumer rights. One of the first steps in the admission procedure for a new candidate variety is a consistent and thorough registration, leading to a description of a number of traits that should uniquely define each variety. Similar trait descriptions are used in other applications like distinctness, uniformity and stability testing (DUS testing). Typical traits relevant for ornamentals are flower color, color distribution and petal shape. For each species the set of traits will differ. This process is time consuming, susceptible to error, and depends on skilled expertise. In this work, we aim to increase the level of automation in this process by using computer vision to estimate/classify the selected traits from images of the flowers, considering real world data sets of roses and gerberas. Using standard deep learning architectures, accuracies of 35-99% have been obtained for selected traits.

1. Introduction

Ornamental flowers have important social, aesthetic, and cultural roles in society, and floriculture is an economically important activity. The Netherlands alone are responsible for more than 50% of the world's production of flowers and ornamental plants with an export value of € 12 billion per year.[1] Of course, it is important to protect the rights of retailers, consumers, and plant breeders developing flower varieties as well as others in the floriculture ecosystem. In this context, a registry or catalog of flower varieties with a set of traits that uniquely define a variety is an important tool. These traits include color or combination of colors, level of filling/opening of the flower center, distribution on the plant (individual or in trusses), and dimensions of the surface, florets, or petals. To name just one application, it can be used to decide if a new candidate variety developed by a breeder is too similar to existing varieties or if it is novel enough to be admitted as a new one. In addition to this aspect of distinctness, similar descriptions are used in uniformity and stability testing (together forming Distinctness Uniformity Stability (DUS) testing).

Currently, this process is done manually by experts who inspect flower samples of a potentially new variety sent by a breeder, and note the respective trait values. For traits that are visible in flower images, this process could be partially automated. Imaging and computer vision have been used in a number of applications in agriculture and the plant sciences such as phenotyping [19,10], plant health monitoring [2], and harvesting [5,17]. In the current context of registering flower characteristics, automation based on image analysis has the potential to be quicker, more consistent (free from subjective interpretation errors by humans), and less dependent on specialized expertise. In addition, automated registration can also be performed on sites where no experts are available, such as *e.g.*, breeders' sites.

Here, we focus on predicting traits for roses and gerberas using deep learning [11,16], with the aim to assess the possibilities of such techniques for characterizing flower varieties. The present paper deals with predicting categorical traits across varieties within species. Color-related traits from the same databases have been addressed in Wehrens et al. [36]. The results presented herein show that it is possible to automate the extraction of certain traits, even though neither the images

* Corresponding author.
E-mail address: manya.afonso@wur.nl (M. Afonso).

[1] https://www.statista.com/topics/3732/flower-industry-in-the-netherlands, accessed in June 2023.

https://doi.org/10.1016/j.atech.2023.100383
Received 17 October 2023; Received in revised form 24 November 2023; Accepted 12 December 2023
Available online 14 December 2023

nor the databases used were set up for this specific purpose. Thus, this initial foray into predicting floricultural traits using machine learning shows promise.

2. Related work

While deep learning and computer vision techniques are already being used in horticultural and agricultural applications, they are usually focused on counting objects (fruits, flowers) via object detection or segmentation, or on identifying the species, and less on obtaining descriptor values for varieties within species, as is the case here. Segmentation and object detection based on classical computer vision or deep learning have been applied to many cases in agriculture, such as tomatoes [3,21], apples [8], and grapes [4] for the purposes of counting, yield prediction, and pruning, respectively.

For flowers, published work usually concentrates on segmentation and counting applications, using deep learning methods such as YOLO [26], FasterRCNN [27] and Deeplab [7]. In Ruizendaal et al. [28], Hemming [13], a FasterRCNN model was used for bounding box detection of gerbera flowers in a production greenhouse, for an automatic harvesting system, with separate object classes for buds and flowers. Another gerbera harvester focused on detecting the stems using classical vision methods [25].

Deep learning architectures for segmentation in 3D of the leaves, stems, and flowers in rose bushes were proposed in Turgut et al. [32, 33], along with a synthetic data set for training [9]. This kind of pointwise segmentation at the level of a plant is necessary for phenotyping as well as robotic operations such as trimming or harvesting [25].

An early work on image level classification was on the species recognition of wild flowers [29]. This approach used hand crafted features such as shape, roundness, number of corners, etc extracted from images of flowers and leaves taken against a black background in laboratory settings, followed by a linear discriminant function classifier. An accuracy of 96% was reported for 34 locally growing species. In Nilsback and Zisserman [20], a support vector machine (SVM) classifier was used with hand-crafted features such as color, shape, texture, boundary shape, spatial distribution of petals, etc leading to a classification accuracy of 73% on the 102-class Oxford flowers data set.

In all these cases machine/deep learning was used to count, segment, or identify flowers. In contrast, the current paper focuses on obtaining descriptors of the objects (flowers) in an automated way.

3. Materials and methods

3.1. Data

We investigate two species, rose and gerbera, each of which is represented by data sets from the two private foundations involved, Floricode[2] and Naktuinbouw.[3] Each of the four data sets comprises a set of images and a database with variety entries, containing the various traits of interest. The images were taken for the purpose of keeping a record, and were thus not intended nor optimized for image analysis. As the two organizations are involved in different tasks, the traits in their respective data sets differ in number and level of detail. Moreover, a trait may have a different meaning depending on the origin of the database, or the flower species. For example, due to the differences in the biological structures of the centers of roses and gerberas, the flower-type of a rose could be either filled or appear as a spinning confluence of the petals. On the other hand for gerberas, the flower-type refers to the level of the center region that is covered by inner ray florets. Here, we focus on categorical traits, having two to five different possible values, depending on the trait.

Clearly, not all traits in the databases could be recognized by considering images alone. Some traits are simply impossible to determine due to characteristics which are invisible in images. Examples include fragance which cannot be captured in an image, but also visible ones that are not captured in the images due to the perspective, such as the presence of thorns in a rose variety. Furthermore, the imaging protocol used in the Floricode dataset leads to flowers being of roughly equal size in the image, regardless of their real size, thus flower sizes and dimensions cannot be determined since the images do not contain any size reference. When prioritizing the categorical traits in the data sets, we first of all focused on those traits that we and the experts from Floricode and Naktuinbouw gave a reasonable probability of identification.

A next consideration concerns the number of examples for each trait and trait category. Not all flower records are completely filled with all traits, and for some traits the number of data points for some classes is simply too small. In addition, since we are looking at categorical traits, all categories of interest should be populated well enough and in a balanced way to allow the deep learning networks to recognize them. We maintain a minimum level of at least 50 examples per category. If a trait has fewer than two categories satisfying this requirement, it is eliminated. In case a trait has several categories that satisfy the count requirement and several others that don't, the number of categories is reduced. If the categories are ordered (*e.g.*, from "very small" to "very large" in, say, five steps) small categories are merged with a neighboring category; if there is no natural ordering, categories that are too small are simply be discarded. These steps lead to the traits for the four data sets listed in Table 1.

3.2. Image pre-processing

Some examples from each data set are shown in Fig. 1. It can be seen that while the images were taken in controlled settings, unlike those of a production greenhouse, there is considerable variation in the background, even within an image in the case of the Naktuinbouw data sets (bottom two rows). For both Floricode data sets, the background is white and relatively uniform. Thus, except for white roses or gerberas, the background can be segmented by applying a threshold on the three image channels to consider as background the pixels for which all three channels are equal and above this threshold. For white flowers, this segmentation is achieved by a combination of color thresholding and edge detection, followed by connected components.

Such a color or edge based approach is not sufficient to segment a non-uniform background, as in the case of the Naktuinbouw data sets. In these cases, the foreground is segmented using the background remover (BGR) app[4] which applies a deep learning model based on the U^2-Net architecture and has been trained on general purpose data sets with semantic foreground background ground truth data [23]. In all cases, the pixels corresponding to the background are set to black (pixel values of zero for each of the R, G and B channels).

Whereas the gerbera data sets and the Naktuinbouw rose data concentrate on the flowers, the Floricode rose data set shows quite a bit of plant leaves. When analyzing the images using deep learning, it appeared that the presence of these leaves did not influence the results, so the images did not receive further pre-processing other than the background removal. This was also true for the color results obtained by deep learning, presented in Wehrens et al. [36], but there it was shown that for a more classical image analysis approach (in this case based on color histograms) the leaves had to be removed before analysis in order to get acceptable results. The fact that deep learning approaches can handle distracting information in the image data is an important advantage in many practical applications.

Some images in these data sets contain top as well as bottom views of two different flower samples of the same variety. In these cases only

[2] www.floricode.com.

[3] www.Naktuinbouw.com.

[4] https://github.com/nadermx/backgroundremover.

Table 1
Characteristics of the four data sets used in this paper.

Dataset	Trait	Classes	# images (train/test)
Floricode, Rose	Flower type	spinning center	195 (136/59)
		filled	318 (222/96)
	Inflorescence	single flowered	1271 (889/382)
		truss	332 (232/100)
Floricode, Gerbera	Flower type	filled	172 (120/52)
		half filled	341 (238/103)
		slightly filled	460 (322/138)
		unfilled	456 (319/137)
	Inflorescence	flattened spherical	76 (53/23)
		little head	1353 (947/406)
Naktuinbouw, Rose	Flower diameter	medium	334 (233/101)
		small	81 (56/25)
	Flower shape	irregularly rounded	413 (289/124)
		round	49 (34/15)
	Petal number	few	149 (104/45)
		many	69 (48/21)
		medium	176 (123/53)
		very many	63 (44/19)
	Petal shape	obovate	380 (266/114)
		transverse elliptic	163 (114/49)
	Petal edge bending	medium	182 (127/55)
		strong	144 (100/44)
		weak	93 (65/28)
	Petal number colors	one	616 (431/185)
		two	173 (121/52)
Naktuinbouw, Gerbera	Disk diameter	large	93 (65/28)
		medium	139 (97/42)
		medium large	97 (67/30)
		small medium	88 (61/27)
		very large	58 (40/18)
	Flower head type	semi double	413 (289/124)
		single	222 (155/67)
	Head diameter	medium	112 (78/34)
		medium large	222 (155/67)
		small medium	246 (172/74)
	Outer floret length	long	182 (127/55)
		medium	163 (114/49)
	Outer floret shape	pointed	171 (119/52)
		rounded	473 (331/142)
	Outer floret width	broad	67 (46/21)
		medium	216 (151/65)

the top view is selected through visual inspection. When the two flower samples in the image have no gap between them, the segmentation is corrected manually by editing the segmented image in a standard image editor.

As a final preprocessing step, the Floricode images (average image size 9 MB) were resized to be similar in size to the Naktuinbouw images (on average 0.9 MB). This not only allows for a more fair comparison, but also speeds up calculations considerably.

3.3. *Deep learning analytics*

For the categorical traits that are the focus of the current paper, the goal of the deep learning methods is to provide a prediction of the trait, given an image of the flower. The prediction problem can be either a binary or a multiple class classification problem, with the labels being the respective trait values. As a neural network architecture we use a standard 18 layer residual network, ResNet18 [12] with an additional layer followed by a softmax classifier over the final fully connected layer, having as many outputs as classes to be predicted. This architecture was previously successfully applied to binary classification problems such as classifying potato plants into healthy or sick [2]. ResNet18, while significantly deeper than architectures such as VGG [30], can work with fewer labeled images than ResNets with more layers. Preliminary calculations with a 50-layer Resnet essentially led to the same results as reported here.

Since some of the categories for the data in this paper have relatively few cases, we decided to split the data into training- and test sets with a division of roughly 70% and 30%, respectively, and to dispense with separate validation sets – these are often used to stop the training process in order to avoid over-fitting. Here, we simply set the number of epochs to a fixed number and stop the training when that number has been reached.

The train-test splits for each trait can be found in Table 1.

3.4. *Software*

The data pre-selection and filtering was done using R [24]. The image pre-processing operations were done using MVTec Halcon[5] [1] for color based foreground segmentation, and python/opencv [6] for edge-based foreground segmentation.

[5] https://www.mvtec.com/products/halcon.

Fig. 1. Examples of raw images from the four data sets. Top to bottom: Floricode rose, Floricode gerbera, Naktuinbouw rose, and Naktuinbouw gerbera. The first and third columns show the original images, and the second and fourth ones show the respective pre-processed images after removing the background.

The deep learning classifiers were implemented using python/ Jupyter notebooks and the PyTorch framework [22]. The final output is a comma separated value (CSV) file containing the true (ground truth) values of the traits and their respective predicted values, over the images in the test set.

4. Results and discussion

To evaluate the performance of our classification models, we focus on results of the test sets, not seen during the training of the deep learning models. Our parameter of interest is accuracy, the proportion of correctly classified samples to the total number of samples. In the case of ordinal variables, an estimate is counted as correct if the predicted category exactly matches the ground truth – a more detailed performance measure could be to calculate distances between predictions and ground truth, where ordered levels are represented by integers, for example. Such a measure would discriminate between gross errors and smaller errors. Since in real-life applications trait databases are going to be used as references, sometimes even with legal implications, here we choose to stay on the conservative side and simply count everything that is not exactly equal to the ground truth as an error. We'll first focus on one trait in particular (flower type for Floricode gerberas) and then present the results for all traits in the four data sets.

4.1. Example trait - Floricode gerbera, flower type

Flower type for Floricode gerberas is an ordered categorical trait, related to the presence and distribution of a set of inner ray florets, also called trans florets. "Unfilled" means that inner ray florets are not present; "slightly-filled" means that there is a thin layer of inner ray florets; "half-filled" means that inner ray florets cover about half of the outer ray florets, whereas "Filled" means that the inner ray florets mostly overlap the outer ray florets. Examples of the four classes are shown in Fig. 2. The differences between the four classes are clearly not immediately obvious; manual classification is hard to do without specific training.

Fig. 3 shows the confusion matrix for the test set of this trait. Although the majority of the cases are predicted correctly (the diagonal is the darkest area of the plot) there are misclassifications, too: the overall accuracy is 72.8%. However, it is clear that most of the errors are found between neighboring classes: especially the "Slightly-filled" class is predicted too often, also in cases where the true label is "Unfilled" or "Half-filled". Apart from the fact that there will be cases which are simply very close to the boundaries of these classes and therefore will be harder to predict, it is also true that the ground-truth labels are not infallible: different observers may arrive at (slightly) different conclusions, especially in a more difficult trait such as this one. In addition, there is some variation between flowers, and the flower that has been used to obtain the database entry may not be the same specimen visible in the image. Finally, real-life databases that have been built up over years, with many different people contributing, are bound to contain some inconsistencies.

Observations such as this does give confidence that the deep-learning models pick up relevant information. However, their black-box nature makes it very hard to find out why a particular answer is given, and where in more classical models it often is possible to assess model coefficients and relate them to expert knowledge, such a common-sense model validation is not straightforward with current interpretation techniques of deep learning. If the model would find the distinction between the two extreme categories equally hard as the difference between neighbors, this could be taken as a warning that maybe the model is focusing on the wrong characteristics (unless all off-

Fig. 2. Training set examples. Data set: Floricode Gerbera, Trait: Flower-type. Class labels (from left to right): (a) Filled, (b) Half-filled, (c) Slightly-filled, (d) Unfilled.

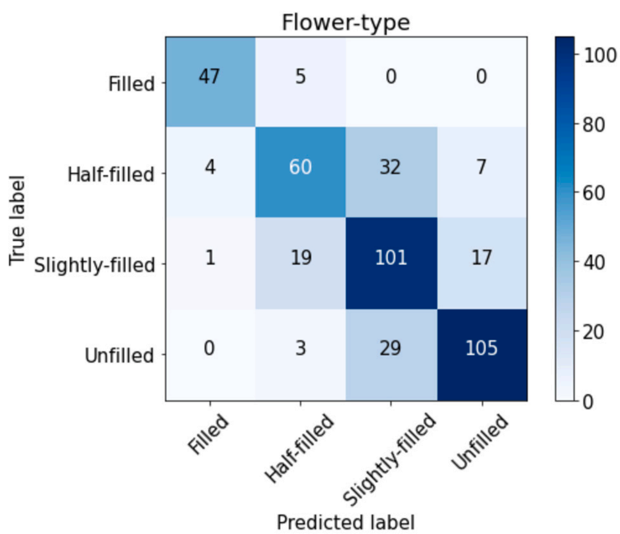

Fig. 3. Prediction results for Floricode gerbera, trait flower type.

diagonal elements in the confusion matrix contain very low numbers, of course). This is clearly not the case here, and the neural networks indeed seem to be able to pick up relevant class information from the images.

4.2. Overall results

In order to assess the variability of the training process, five repeated training runs with the Resnet18 architecture were done starting from different (random) initializations. The average classification error rates for all the traits, again based on predictions of unseen test sets, are shown in Fig. 4. The figure also shows the 95% confidence intervals over the multiple runs per trait. Some spread is visible, but it is clear that the differences in performance across traits cannot be attributed to random learning effects alone.

Mean accuracies range from 35.2% (Naktuinbouw Disk-diameter) to 99.1% (Floricode Inflorescence). This large difference is caused by several factors. First of all, some traits have more clear cut differences between the classes and are thus easier than other. An example is the rose inflorescence in the Floricode data set, *e.g.*, which is the difference between plants where roses are individual flowers or trusses of flowers, something that is very easily seen in the images. Other traits, such as the Floricode Flower-type trait for gerberas as discussed in Section 4.1, are much more difficult.

Secondly, traits with multiple possible values are much more difficult to predict than traits with only two possible values (binary classification problems). In the Naktuinbouw data sets this is clearly visible: the two worst traits in roses are the ones with four (Petal number) and three (Petal edge bending) levels – the worst trait overall (Naktuinbouw gerbera, Disk diameter) even has five categories. Again, we are only counting exact matches with the ground truth as correct predictions. Inspecting the confusion matrices for all ordered traits, however, in all cases we see the behavior also observed for Floricode gerbera Flower type: the confusion is mainly between neighboring classes, and much less between classes that are far apart.

Thirdly, imbalanced class sizes are difficult for prediction models. The best score, *e.g.*, is obtained for Floricode gerbera Inflorescence, but there the test set contains only 23 examples for “flattened spherical”, and 406 examples for the other category, “little head”. In other words, predicting everything as “little head” would already lead to a 94.6% accuracy. The very fact that the model is able to improve on this number is surely impressive, but the context is needed to view the accuracy in the correct light. Note that the imbalance is not only a difficulty in training the network, but also in the evaluation of the results.

Overall, it is clear that the images provide information to learn different traits, and although most of the accuracies should still be increased before they are useful in everyday practice this work shows definite promise. Several ways are available to achieve this – tuning the network architecture and hyper-parameters to the trait at hand will definitely help. Heat-map analysis could also be worth investigating by way of explainability, to determine what the network learns. Data augmentation by translation and/or rotation could be helpful in the presence of contextual information such as stems or leaves.

5. Conclusions and future work

In this work, we have reported results on a first foray of deep learning into the estimation of floricultural traits. Reasonable results were obtained even though the data have been acquired over a number of years, with different observers, and not specifically for this objective.

Several avenues are available to further improve results for the current set of traits. The first is simply to increase data set sizes, perhaps by merging databases from growers or other (international) organizations using the same sets of characteristics. This not only increases coverage for the less populated classes, but in general will lead to better, more accurate and more robust AI models. The second is to pursue an even higher level of data standardization. On the image side of things, this could be achieved by controlled lighting conditions and image acquisition protocols – sometimes very simple elements like including color

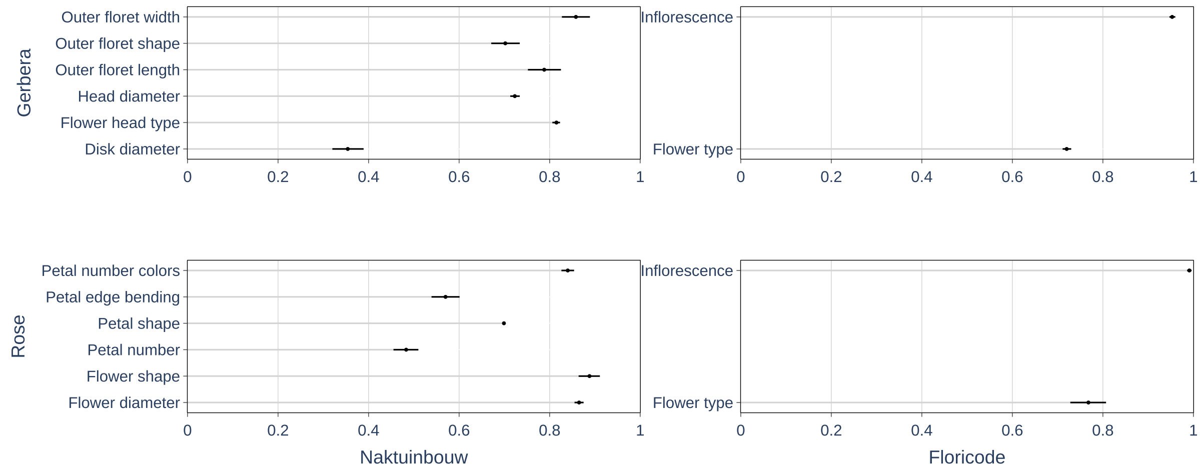

Fig. 4. Average classification accuracies for the traits, grouped by data set, based on five repeated runs. 95% Confidence intervals, based on the spread over these runs, are indicated in the figure by horizontal line segments.

reference charts or rulers to measure absolute sizes in the images will make already a big difference. Furthermore, more advanced learning architectures can be assessed and optimized for specific traits. For traits such as sizes or lengths, regression models [37,31] could have better performance with data comprising actual measured dimensions. Ordinal regression [34] in which ordered categorical classes are predicted may also be a helpful approach. CNNs with attention modules [38] which work with soft embeddings rather than hard labels, as well as attention followed by text decoder model for generating description of visual traits of flowers [18,14] may be approaches worth investigating. The latter will require training on a broader dataset but specifically related to plants. Finally it might be interesting to investigate transfer learning of similar traits across flower species, *i.e.*, if models trained on one species or dataset can be applied on a different species or dataset.

Synthetic data obtained through methods such as Generative Adversarial Networks (GANs) are also worth investigating as a means of generating training data [35]. Extending the deep learning approach to object detection with different classes could offer a way to detect flowers with defects or blemishes.

Conversely, the question could in some cases be turned around: rather than investigating which traits of a predefined set can be obtained from image material, we can ask ourselves what information can be retrieved from the images – this could form the basis of a new set of characteristics, leading to a different but perhaps equally valid description of ornamentals, most probably being complemented with other descriptors based on genetics. While this paper addressed only those traits with information present in 2D images, using multiple cameras or RGB-plus-depth sensors such as Intel Realsense cameras [15] could be used to reconstruct a point cloud representation and extract information on 3D traits such as cross-sectional profile and petal structure.

Automation of the process also makes it possible to decentralize the trait assessment, and registration could become similar to sending in a digital tax form. The reproducibility of the results based on generally available models, would form a solid legal basis for such a procedure.

CRediT authorship contribution statement

Manya Afonso: Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Maria-João Paulo:** Data curation, Formal analysis, Investigation, Writing – review & editing. **Hubert Fonteijn:** Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. **Mary van den Helder:** Data curation, Writing – review & editing. **Henk Zwinkels:** Data curation, Funding acquisition, Project administration, Writing – review & editing. **Marcel Rijsbergen:** Data curation, Writing – review & editing. **Gerard van Hameren:** Data curation, Writing – review & editing. **Raoul Haegens:** Data curation, Funding acquisition, Project administration, Writing – review & editing. **Ron Wehrens:** Conceptualization, Formal analysis, Funding acquisition, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The data that has been used is confidential.

Acknowledgements

This work was supported by co-funding from foundation TKI (Topsector Knowledge and Innovation), Horticulture and Starting Materials, grant number LWV20.172.

References

[1] Halcon user guide, https://www.mvtec.com/fileadmin/Redaktion/mvtec.com/products/halcon/documentation/manuals/hdevelop_users_guide.pdf. (Accessed 28 March 2023), 2022.

[2] M. Afonso, P.M. Blok, G. Polder, J.M. van der Wolf, J. Kamp, Blackleg detection in potato plants using convolutional neural networks, in: 6th IFAC Conference on Sensing, Control and Automation Technologies for Agriculture AGRICONTROL 2019, IFAC-PapersOnLine (ISSN 2405-8963) 52 (30) (2019) 6–11, https://doi.org/10.1016/j.ifacol.2019.12.481.

[3] M. Afonso, A. Mencarelli, G. Polder, R. Wehrens, D. Lensink, N. Faber, Detection of tomato flowers from greenhouse images using colorspace transformations, in: P.M. Oliveira, P. Novais, L.P. Reis (Eds.), EPIA 2019, in: Lecture Notes in Artificial Intelligence, vol. 11804, Springer Nature, Switzerland AG, 2019, pp. 146–155.

[4] A. Aquino, B. Millan, S. Gutiérrez, J. Tardáguila, Grapevine flower estimation by applying artificial vision techniques on images with uncontrolled scene and multi-model analysis, Comput. Electron. Agric. (ISSN 0168-1699) 119 (2015) 92–104, https://doi.org/10.1016/j.compag.2015.10.009.

[5] B. Arad, J. Balendonck, R. Barth, O. Ben-Shahar, Y. Edan, T. Hellström, J. Hemming, P. Kurtser, O. Ringdahl, T. Tielen, et al., Development of a sweet pepper harvesting robot, J. Field Robot. 37 (6) (2020) 1027–1039.

[6] G. Bradski, The opencv library, Dr. Dobb's J. Softw. Tools Prof. Program. 25 (11) (2000) 120–123.

[7] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs, IEEE Trans. Pattern Anal. Mach. Intell. (ISSN 0162-8828) 40 (4) (April 2018) 834–848, https://doi.org/10.1109/TPAMI.2017.2699184.

[8] P.A. Dias, A. Tabb, H. Medeiros, Apple flower detection using deep convolutional networks, Comput. Ind. (ISSN 0166-3615) 99 (2018) 17–28, https://doi.org/10.1016/j.compind.2018.03.010.

[9] H. Dutagaci, P. Rasti, G. Galopin, D. Rousseau, ROSE-X: an annotated data set for evaluation of 3D plant organ segmentation methods, Plant Methods 16 (1) (2020) 1–14.

[10] H. Fonteijn, M. Afonso, M. Mooij, N. Faber, A. Vroegop, G. Polder, R. Wehrens, Automatic phenotyping of tomatoes in production greenhouses using robotics and computer vision: from theory to practice, Agronomy 11 (8) (2021) 1599.

[11] I. Goodfellow, Y. Bengio, A. Courville, Deep Learning, MIT Press, 2016.

[12] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 770–778.

[13] J. Hemming, Current developments in greenhouse robotics and challenges for the future, Acta Hortic. 1296 (2020) 975–985.

[14] M.D. Zakir Hossain, Ferdous Sohel, Mohd Fairuz Shiratuddin, Hamid Laga, A comprehensive survey of deep learning for image captioning, ACM Comput. Surv. 51 (6) (2019) 1–36.

[15] L. Keselman, J. Iselin Woodfill, A. Grunnet-Jepsen, A. Bhowmik, Intel realsense stereoscopic depth cameras, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2017, pp. 1–10.

[16] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436.

[17] C.F. Lehnert, A. English, C. McCool, A.W. Tow, T. Perez, Autonomous sweet pepper harvesting for protected cropping systems, IEEE Robot. Autom. Lett. 2 (2) (2017) 872–879.

[18] Zhihong Lin, Donghao Zhang, Qingyi Tao, Danli Shi, Gholamreza Haffari, Qi Wu, Mingguang He, Zongyuan Ge, Medical visual question answering: a survey, Artif. Intell. Med. (2023) 102611.

[19] M. Minervini, H. Scharr, S.A. Tsaftaris, Image analysis: the new bottleneck in plant phenotyping [applications corner], IEEE Signal Process. Mag. (ISSN 1558-0792) 32 (4) (July 2015) 126–131, https://doi.org/10.1109/MSP.2015.2405111.

[20] M.E. Nilsback, A. Zisserman, Automated flower classification over a large number of classes, in: 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, IEEE, 2008, pp. 722–729.

[21] D. Oppenheim, Y. Edan, G. Shani, Detecting tomato flowers in greenhouses using computer vision, Int. J. Comput. Electr. Autom. Control Inf. Eng. 11 (1) (2017) 104–109, World Academy of Science, Engineering and Technology.

[22] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Kopf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, S. Chintala, Pytorch: An Imperative Style, High-Performance Deep Learning Library, Advances in Neural Information Processing Systems, vol. 32, Curran Associates, Inc., 2019, pp. 8024–8035, http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf.

[23] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O.R. Zaiane, M. Jagersand, U2-net: going deeper with nested u-structure for salient object detection, Pattern Recognit. 106 (2020) 107404.

[24] R Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, 2022, https://www.R-project.org/.

[25] T. Rath, M. Kawollek, Robotic harvesting of Gerbera Jamesonii based on detection and three-dimensional modeling of cut flower pedicels, Comput. Electron. Agric. 66 (2009) 85, https://doi.org/10.1016/j.compag.2008.12.006.

[26] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, in: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2016, pp. 779–788.

[27] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, in: Advances in Neural Information Processing Systems, 2015, pp. 91–99.

[28] J. Ruizendaal, R. Barth, J. Balendonk, Flower and bud detection of Gerbera L. using faster region-based convolutional networks, in: Proceedings of the International Conference of Agricultural Engineering, 2018.

[29] T. Saitoh, T. Kaneko, Automatic recognition of wild flowers, in: Proceedings - International Conference on Pattern Recognition, vol. 2, 02 2000, pp. 507–510.

[30] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint, arXiv:1409.1556, 2014.

[31] Nima Teimouri, Mads Dyrmann, Per Rydahl Nielsen, Solvejg Kopp Mathiassen, Gayle J. Somerville, Rasmus Nyholm Jørgensen, Weed growth stage estimator using deep convolutional neural networks, Sensors (ISSN 1424-8220) 18 (5) (2018), https://doi.org/10.3390/s18051580, https://www.mdpi.com/1424-8220/18/5/1580.

[32] K. Turgut, H. Dutagaci, G. Galopin, D. Rousseau, Segmentation of structural parts of rosebush plants with 3D point-based deep learning methods, Plant Methods 18 (1) (2022) 20.

[33] K. Turgut, H. Dutagaci, D. Rousseau, Rosesegnet: an attention-based deep learning architecture for organ segmentation of plants, Biosyst. Eng. 221 (2022) 138–153.

[34] Gerhard Tutz, Ordinal regression: a review and a taxonomy of models, Wiley Interdiscip. Rev.: Comput. Stat. 14 (2) (2022) e1545.

[35] K. Wang, C. Gou, Y. Duan, Y. Lin, X. Zheng, F.Y. Wang, Generative adversarial networks: introduction and outlook, IEEE/CAA J. Autom. Sin. 4 (4) (2017) 588–598.

[36] R. Wehrens, M. Afonso, H. Fonteijn, J. Paulo, G. Polder, H. Zwinkels, M. Rijsbergen, G. van Hameren, R. Haegens, Determining flower colors from images, Euphytica (2023), https://doi.org/10.1007/s10681-023-03258-2.

[37] Lingxian Zhang, Zanyu Xu, Dan Xu, Juncheng Ma, Yingyi Chen, Zetian Fu, Growth monitoring of greenhouse lettuce based on a convolutional neural network, Hortic. Res. 7 (1) (2020) 124.

[38] Menghua Zheng, Jiayu Xu, Yinjie Shen, Chunwei Tian, Jian Li, Lunke Fei, Ming Zong, Xiaoyang Liu, Attention-based cnns for image classification: a survey, J. Phys. Conf. Ser. 2171 (1) (jan 2022) 012068, https://doi.org/10.1088/1742-6596/2171/1/012068.