ORIGINAL RESEARCH **Open access**

AI-Based Trait Estimation for Ornamentals in Variety Registration Applications

Yifan Zhao[1*], Liang Chen[1], Ahmad Fauzi[2], Nor Hayati[3]

Abstract

Safeguarding the intellectual property of plant breeders and the entitlements of consumers relies critically on the official recording of novel ornamental flower varieties. A fundamental early stage in the acceptance protocol for a new candidate variety is a rigorous, uniform registration procedure that yields a profile of characteristics intended to serve as a unique identifier for each variety. Equivalent characterizations are employed in other settings, including distinctness, uniformity, and stability (DUS) testing. For ornamentals, typical traits of interest comprise flower color, the patterning of that color, and petal morphology, with the precise suite of traits depending on the species. This workflow requires significant time, is prone to inaccuracies, and relies on specialist knowledge. This investigation aims to increase automation in this workflow by leveraging computer vision to estimate and classify targeted traits from floral images, drawing on real-world datasets of roses and gerberas. Accuracy levels ranging from 35% to 99% have been achieved for selected traits using standard deep learning architectures.

Keywords Deep learning, Computer vision, Floriculture, Ornamentals, Variety registration

*Correspondence:
Yifan Zhao
yifan.zhao@cau.edu.cn
[1] Department of AI for Rice-Fish and Rice-Duck Integrated Farming Systems, School of Agriculture, China Agricultural University, Beijing, China
[2] Department of Smart Aquatic Weed and Algae Prediction, Faculty of Agriculture, University of Malaya, Kuala Lumpur, Malaysia
[3] Department of Machine Learning for Water Quality Monitoring, Faculty of Agriculture, Universiti Putra Malaysia, Serdang, Malaysia

Introduction

Within society, ornamental flowers fulfill crucial social, aesthetic, and cultural functions, and the floriculture industry carries substantial economic weight. Over half of the worldwide production of flowers and ornamental plants originates from the Netherlands alone, generating an annual export value of 12 billion [1]. It is vital to defend the interests of retailers, consumers, breeders developing flower varieties, and other stakeholders in the floriculture ecosystem. In this regard, a registry or catalog documenting varieties based on a set of uniquely defining traits is a key instrument. Such traits can include color or color blends, the extent of the flower center's fill or openness, arrangement on the plant (solitary or in clusters), and measurements of the bloom, florets, or petals. To give but one use case, it serves to determine whether a novel candidate variety bred by a developer is excessively similar to pre-existing varieties or sufficiently novel to warrant official recognition. Beyond this test of distinctness, analogous descriptions form the basis for uniformity and stability evaluations (jointly termed Distinctness, Uniformity, and Stability (DUS) testing).

Today, this task is carried out manually by skilled specialists who inspect flower specimens of a potentially new variety submitted by a breeder and log the corresponding trait assessments. When traits are observable in flower images, there is an opportunity to automate this procedure partially. Imaging and computer vision technologies have found numerous roles in agriculture and plant sciences, including phenotyping [2, 3], plant health surveillance [4], and harvesting [5, 6]. In the specific domain of documenting flower characteristics, image analysis-based automation offers the potential for

This is an open access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License (CC BY-NC-SA 4.0), which permits non-commercial use, distribution, reproduction, and adaptation in any medium or format, provided that appropriate credit is given to the original author(s) and source, a link to the license is provided, and any changes made are indicated. Adapted material must be distributed under the same license.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

greater speed, enhanced reproducibility (by eliminating subjective human judgment errors), and reduced dependency on specialized training. Automated registration could also be performed on-site at locations lacking expert personnel, such as breeding facilities.

Here, our investigation centers on forecasting traits for roses and gerberas using deep learning methodologies [7, 8] to assess the viability of these techniques for cataloging flower varieties. The current study tackles the prediction of categorical traits across multiple varieties within a species. Traits related to color from the same data repositories have been explored separately by Wehrens *et al.* [9]. The outcomes presented herein confirm the feasibility of automating the extraction of specific traits, despite neither the imagery nor the databases being purpose-built for this task. Consequently, this pilot exploration into machine learning-based prediction of floricultural traits indicates considerable promise.

Related Work

While deep learning and computer vision are already deployed in horticultural and agricultural contexts, their application typically targets the quantification of objects (fruits, flowers) via detection or segmentation networks, or species-level classification, rather than the retrieval of descriptor values for distinct varieties within a species, which is the focus here. Segmentation and object detection, using either classical computer vision or deep learning, have been applied to numerous agricultural scenarios, such as tomatoes [10, 11], apples [12], and grapes [13] for purposes of enumeration, yield forecasting, and pruning, respectively.

Published work on flowers predominantly addresses segmentation and counting tasks, using deep learning frameworks such as YOLO [14], Faster R-CNN [15], and Deeplab [16]. Studies by Ruizendaal *et al.* [17] and Hemming [18] used a Faster R-CNN model for bounding-box detection of gerbera blooms in a commercial greenhouse, aimed at an autonomous harvesting platform, distinguishing between object categories for buds and open flowers. A separate gerbera harvesting initiative employed classical vision approaches to detect stems [19].

Three-dimensional deep learning architectures for segmenting leaves, stems, and flowers in rose bushes were introduced by Turgut *et al.* [20, 21], along with a synthetic training dataset [22]. Achieving point-wise segmentation at the whole-plant scale of this nature is a prerequisite for phenotyping activities and robotic interventions like pruning or harvesting [19].

A pioneering image-level classification study addressed the recognition of wild flower species [23]. This method relied on hand-engineered features, such as shape characteristics, roundness, and corner counts, derived from laboratory images of flowers and leaves set against a black backdrop, followed by a linear discriminant function classifier. The reported accuracy reached 96% across 34 locally occurring species. In the work of Nilsback and Zisserman [24], a support vector machine (SVM) classifier was paired with hand-crafted features covering color, shape, texture, boundary contour, and the spatial layout of petals, achieving a classification accuracy of 73% on the 102-class Oxford flowers dataset.

Across all these cases, machine or deep learning served to count, segment, or identify flowers. In stark contrast, the present paper focuses on the automated derivation of descriptors for the objects themselves (flowers).

Materials and Methods

Data

Two species form the basis of our study, rose and gerbera, each represented by data contributions from the two partner private foundations, Floricode[2] and Naktuinbouw.[3] Each of the four resulting datasets consists of an image collection and a database containing a variety of records populated with the different traits under investigation. The photographs were originally shot for archival documentation and were thus neither collected with nor optimized for subsequent image analysis. Because the two organizations engage in distinct functions, the traits cataloged in their respective databases vary in quantity and granularity. Additionally, a trait's interpretation can shift depending on the originating database or the

© 2024 The Author(s). | Published by Agro Intelligence Research Press

floral species involved. To illustrate, owing to the divergent biological architecture of rose and gerbera centers, the flower-type designation for a rose might be “filled” or manifest as a spiraling convergence of petals. In contrast, for gerberas, the same term describes the degree to which inner ray florets obscure the central disc. Our analysis focuses on categorical traits with 2-5 possible states, depending on the trait.

Not every trait within the databases could be discerned from images alone. Certain traits are fundamentally undeterminable because the relevant characteristics are not visible in photographs. Examples range from fragrance, which is inherently non-visual, to visible features omitted from the imagery due to the chosen camera angle, such as thorns on a rose stem. Furthermore, the imaging protocol applied in the Floricode dataset standardizes flowers to roughly uniform sizes in the photographs, irrespective of their true dimensions; consequently, bloom sizes and measurements are impossible to ascertain, as the images include no scale reference. In ranking the categorical traits by priority, we concentrated foremost on those that both our team and the specialists at Floricode and Naktuinbouw considered to have a reasonable likelihood of successful identification.

An additional factor was the number of available examples for each trait and its constituent categories. Not every variety record is fully populated across all traits, and for certain traits, the volume of data points for particular classes is simply inadequate. Since the focus is on categorical traits, all categories of concern must be sufficiently and evenly represented to permit the deep learning networks to learn their distinguishing features. We instituted a minimum of 50 examples per category. Any trait possessing fewer than two categories that meet this threshold is removed from consideration. If a trait includes several categories that satisfy the numerical requirement alongside others that fall short, the total number of categories is consolidated, where the categories follow a logical sequence (e.g., from “very small” to “very large” over five tiers), sparsely populated categories are folded into an adjacent one; where no natural progression exists, undersized categories are discarded. This filtering procedure resulted in the final selection of traits for the four datasets, as presented in **Table 1**.

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)

© 2024 The Author(s). | Published by Agro Intelligence Research Press

		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)

Image pre-processing

A selection of specimens from each dataset is illustrated in **Figure 1**. One can observe that, notwithstanding the controlled capture environment—unlike the conditions of a commercial greenhouse—substantial background heterogeneity exists, including within single photographs for the Naktuinbouw collections (lower pair of rows). The two Floricode collections feature a white, relatively homogeneous backdrop. Hence, barring instances of white roses or gerberas, the backdrop can be isolated by establishing a threshold across the trio of image channels, classifying any pixel

© 2024 The Author(s). | Published by Agro Intelligence Research Press

as background for which all three channels hold equal values above this cutoff. For white blossoms, separation is achieved by combining color thresholding with edge detection, followed by a connected components routine.



Figure 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 columns show the pre-processed images after background removal.

A strategy based on color or edges lacks the sophistication needed to partition a non-uniform setting, as characterized by the Naktuinbouw collections. Under these circumstances, the foreground is delineated using the background remover (BGR) application[4], which leverages a deep learning model based on the U-Net architecture, previously trained on general-purpose datasets with semantic foreground-background ground truth labeling [25]. In general, the background pixels are forced to black (zero intensity in each R, G, and B channel).

Where the Gerbera collections and the Naktuinbouw rose data keep the blooms central, the Floricode rose collection features abundant foliage. Upon processing the imagery with deep learning, the presence of these leaves was found not to influence the outcomes; accordingly, the images were spared any additional pre-treatment beyond background stripping. This finding was corroborated by the color-focused deep learning results reported in Wehrens *et al.* [9]. Yet, that

© 2024 The Author(s). | Published by Agro Intelligence Research Press

same study revealed that a more conventional image analysis pipeline (specifically one anchored in color histograms) required leaf excision before analysis to yield usable outcomes. The ability of deep learning methodologies to ignore extraneous signals in image data is a significant asset across many real-world use cases.

Several pictures across these collections simultaneously capture top and bottom perspectives of two separate flower specimens belonging to a single variety. Where this occurs, solely the overhead view is preserved via manual visual triage. For pictures in which the pair of flower specimens touch without any intervening space, the resulting segmentation is subsequently corrected by hand using a common image manipulation program.

In a terminal pre-processing action, the Floricode pictures (mean file size of 9 MB) were reduced in resolution to align more closely with the dimensions of the Naktuinbouw pictures (mean file size of 0.9 MB). This step serves the dual purpose of enabling a more balanced comparison and markedly shortening computation times.

Deep learning analytics

Concerning the categorical traits that constitute the core subject of this article, the deep learning system aims to output a trait estimate conditioned on an input flower image. The estimation challenge may take the form of a binary distinction or a multi-way classification, with the possible trait values acting as the target labels. The chosen network backbone is a conventional 18-layer residual architecture, ResNet18 [26], fitted with an extra layer and a softmax classifier stacked on top of the final fully connected layer, with the number of outputs matching the number of target classes. This design has formerly proven effective for binary discrimination problems, exemplified by the sorting of potato plants into healthy and diseased categories [4]. ResNet18, despite being considerably deeper than predecessors like VGG [27], remains trainable with fewer annotated samples than deeper ResNet variants. Exploratory runs with a 50-layer ResNet yielded outcomes essentially indistinguishable from those documented here.

Since several categories within this investigation's data have comparatively limited sample counts, we divided the material into training and test partitions at an approximate 70% to 30% ratio, electing to omit a dedicated validation partition—such partitions are commonly employed to interrupt training and thereby counteract overfitting. In this work, we set the epoch count to a predetermined value and cease training upon its attainment.

The train-test partitioning for each trait is enumerated in **Table 1**.

Software

Initial data curation and filtering were conducted within the R environment [28]. Image pre-processing tasks were executed using MVTec Halcon[5] [1] for color-based foreground separation and Python/OpenCV [29] for edge-based foreground separation.

The deep learning classifiers were built using Python/Jupyter notebooks atop the PyTorch library [30]. The ultimate product is a comma-separated values (CSV) file that stores the authentic (ground-truth) trait designations alongside their model-inferred counterparts for all images in the test partition.

Results and Discussion

When appraising the efficacy of our classification pipelines, we focus on outcomes in the test partitions, deliberately withheld during the deep learning models' training phase. The quantity under examination is accuracy: the fraction of correctly assigned samples relative to the total number of samples. With ordinal variables, a prediction is scored as correct only when the inferred category matches the ground truth exactly—a richer performance metric could involve computing the distance between predictions and ground truth, perhaps by encoding ordered tiers as numerical values. A scheme of this sort would distinguish major blunders from marginal ones. Bearing in mind that, within practical

© 2024 The Author(s). | Published by Agro Intelligence Research Press

deployments, trait registries are destined to function as authoritative references, occasionally carrying legal consequences, we deliberately adopt a cautious philosophy here and deem any inference that falls short of an exact ground truth correspondence as erroneous. Our examination commences with a deep dive into a representative trait (flower type within the Floricode gerbera set), followed by a summary of the collective findings across all traits spanning the four collections.

Example trait - Floricode gerbera, flower type

Flower type for Floricode gerberas is an ordered categorical descriptor associated with the existence and spatial spread of a whorl of interior ray florets, alternatively referred to as trans florets. “Unfilled” conveys that internal ray florets are absent; “slightly-filled” conveys the existence of a sparse ring of such florets; “half-filled” conveys that the internal ray florets obscure roughly fifty percent of the exterior ray florets, whereas “Filled” conveys that the internal ray florets substantially overshadow the exterior ray florets. Representative specimens of the four groupings appear in **Figure 2**. The contrasts separating these four groupings are manifestly not self-evident upon casual inspection; manual sorting proves demanding in the absence of targeted instruction.



Figure 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, and (d) Unfilled.

Figure 3 reproduces the confusion matrix computed over the test partition for this particular trait. Although the preponderance of observations receives a correct assignment (the matrix diagonal stands as the most intensely shaded zone), misattributions are likewise present: the aggregate accuracy reaches 72.8%. Even so, one may plainly see that mistaken assignments occur overwhelmingly between contiguous classes: the “Slightly-filled” designation, most notably, suffers from over-prediction, surfacing even where the authentic label reads “Unfilled” or “Half-filled”. Quite apart from the reality that a subset of observations will inherently reside extremely close to the frontiers dividing these classes and will therefore present heightened predictive difficulty, it remains the case that the ground-truth annotations themselves are not immune to error: separate evaluators can arrive at (marginally) differing judgments, particularly when confronting a more demanding trait such as this one. Moreover, there is biological variation among individual flowers, and the precise bloom used to populate the database record may differ from the one shown in the photograph. To conclude, operational databases assembled over extended timeframes, shaped by the input of a rotating cast of contributors, are virtually guaranteed to harbor some measure of internal inconsistency.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

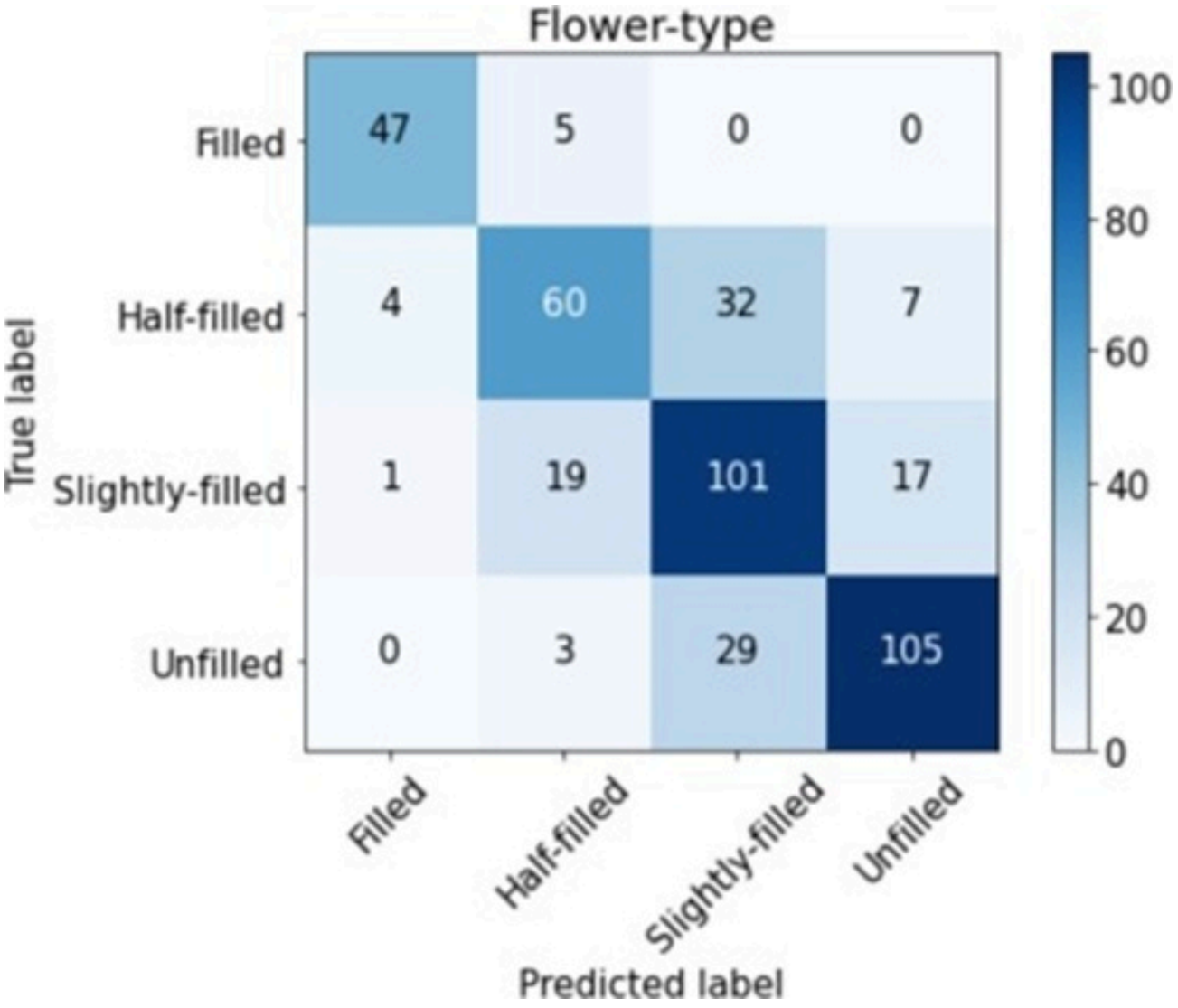

Figure 3. Prediction results for Floricode gerbera, trait flower type.

Insights of this kind foster the trust that the deep-learning machinery is registering a meaningful signal. That being said, the opaque, black-box constitution of such models renders it exceedingly challenging to unpack the rationale behind any single output, and while in the realm of more traditional modeling, it frequently remains feasible to scrutinize model parameters and map them onto domain expertise, a comparable form of intuitive model vetting does not come naturally to today's deep learning explanation toolkits. Were the model to find differentiating between the pair of polar categories equally taxing as differentiating between adjoining categories, this pattern could be interpreted as a red flag suggesting the model might be latching onto spurious features (setting aside, naturally, scenarios where every off-diagonal cell within the confusion matrix registers negligible counts). Such is plainly not the scenario here, and the neural architectures indeed seem well able to harvest meaningful class-discriminative cues from the visual data.

Overall results

To gauge the stochastic fluctuation inherent in the training regimen, five independent training cycles were performed using the ResNet18 design, each initialized from a distinct (random) starting point. The mean classification error percentages aggregated across all traits, again based on inferences from unseen test partitions, are shown in **Figure 4**. The figure also displays the 95% confidence intervals computed across replicate runs for each trait. A certain amount of

© 2024 The Author(s). | Published by Agro Intelligence Research Press

variance is evident, yet it is unmistakable that the differentials in trait-specific performance cannot be chalked up exclusively to stochastic learning phenomena.

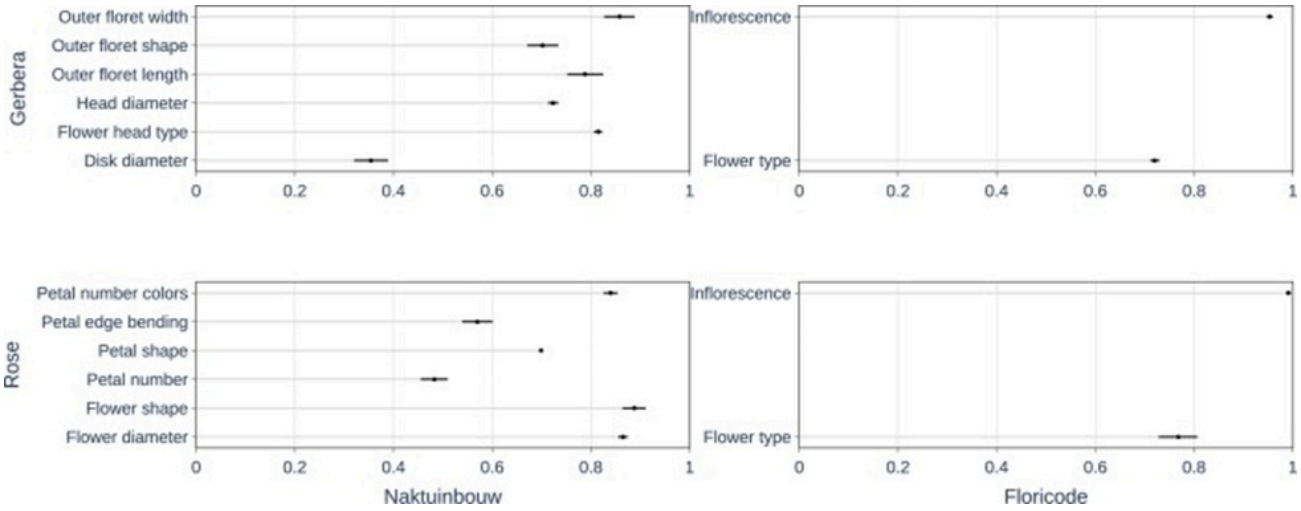

Figure 4. Average classification accuracies for the traits, grouped by data set, based on five repeated runs. 95% Confidence intervals, based on the spread over these runs, are indicated in the figure by horizontal line segments.

Mean accuracies stretch from a low of 35.2% (Naktuinbouw Disk-diameter) to a high of 99.1% (Floricode Inflorescence). This expansive spread can be traced to multiple causative elements. To begin with, a subset of traits exhibits sharper inter-class demarcations and consequently presents a simpler learning problem. A case in point is the rose inflorescence trait within the Floricode collection, which captures the divergence between plants bearing isolated flowers versus those bearing flower clusters, a visual feature trivially discerned in photographs. Other traits, epitomized by the Floricode Flower-type descriptor for gerberas elaborated in Section 4.1, pose far steeper challenges.

In the second place, descriptors with a multiplicity of admissible values prove markedly more resistant to accurate prediction than those constrained to merely two alternatives (binary classification settings). This pattern is unmistakable within the Naktuinbouw collections: the pair of weakest-performing descriptors in roses are precisely those harboring four (Petal number) and three (Petal edge bending) tiers—the single weakest descriptor across the board (Naktuinbouw gerbera, Disk diameter) even extends to five categories. It bears reiterating that we count exclusively exact correspondences with the ground truth as successful predictions. Perusing the confusion matrices for every ordered trait, nonetheless, confirms in each instance the behavioral tendency already witnessed for Floricode gerbera Flower type: misattributions are largely confined to neighboring classes, with far rarer confusion arising between well-separated classes.

In the third place, lopsided class distributions pose obstacles for predictive systems. The topmost accuracy score, by way of illustration, is registered for Floricode gerbera Inflorescence, where the test partition for that trait houses a mere 23 instances of “flattened spherical” set against 406 instances of the alternative category, “little head”. To put it differently, blindly guessing “little head” for every sample would furnish a 94.6% accuracy straight away. The sheer fact that the model outstrips this naïve baseline is undoubtedly noteworthy, yet this contextual backdrop is indispensable for interpreting the accuracy figure appropriately. One should note that the imbalance poses a hurdle not only for network optimization but also for subsequent performance evaluation.

Taken in sum, it is apparent that the imagery encodes sufficient information for models to assimilate a range of traits, and although the majority of accuracy figures would still require improvement before attaining suitability for routine operational deployment, the present study showcases unambiguous promise. Multiple levers exist to realize such enhancements—customizing the network topology and hyperparameter selections to the idiosyncrasies of each given trait is certain to

© 2024 The Author(s). | Published by Agro Intelligence Research Press

yield dividends. The application of heat-map visualization also stands out as a worthwhile avenue to pursue in the name of interpretability, enabling precise pinpointing of what the network is internalizing. Data augmentation strategies involving translation and/or rotation might confer advantages in situations where contextual elements like stems or foliage are present.

Conclusion

In this contribution, we have chronicled the findings of a preliminary study on the deployment of deep learning for deducing floricultural descriptors. Encouraging results were achieved despite the data having been gathered over years by a shifting roster of evaluators, and without being expressly commissioned for this specific objective.

Numerous trajectories remain open further to improve outcomes for the existing suite of descriptors. The most immediate is the straightforward expansion of dataset volumes, perhaps achieved by fusing registries maintained by cultivators or by additional (international) bodies that adhere to identical descriptor frameworks. Doing so would not only bolster representation for sparsely populated classes but would also broadly yield AI models of superior accuracy, robustness, and reliability. A second trajectory entails striving for an even loftier tier of data uniformity. On the visual side of the equation, this ambition could be fulfilled by instituting regulated illumination setups and standardized photographic procedures—on occasion, remarkably elementary additions, such as the inclusion of color calibration targets or dimensional scales for deriving absolute measurements from photographs, would alone constitute a meaningful leap forward. Moving beyond that, more elaborate learning architectures can be evaluated and fine-tuned for individual descriptors. For descriptors of a mensurational nature, such as sizes or lengths, regression-oriented models [31, 32] could deliver improved performance when paired with datasets that include actual measured quantities. Ordinal regression frameworks [33], engineered to forecast ordered categorical tiers, could also represent a fruitful methodological direction. Convolutional neural networks augmented with attention modules [34] that engage with soft embeddings rather than rigid labels, together with attention mechanisms feeding into text-decoder components for the automatic composition of visual trait descriptions for flowers [35, 36], constitute additional approaches worthy of exploration. The latter approach, however, would demand training on a more expansive, botanically focused corpus. As a final consideration, exploring transfer learning, in which models trained on one species or data environment are redeployed to a distinct one, may yield rewarding insights.

Artificially generated data, fabricated via techniques such as Generative Adversarial Networks (GANs), similarly merit examination as a supplementary source of training exemplars [37]. Broadening the deep learning paradigm to encompass object detection with differentiated class labels could supply a mechanism for flagging blooms exhibiting imperfections or lesions.

From the reverse angle, there are instances where the framing of the inquiry could be flipped. Instead of asking which descriptors from a preordained set can be extracted from image data, we might ask what information the images can yield. The answer could lay the groundwork for an entirely new descriptor vocabulary, eventuating in an alternative, though conceivably no less legitimate, schema for characterizing ornamentals—one that would in all probability be supplemented by further descriptors anchored in genetic profiling. While this study restricted itself to descriptors whose informational content is present within two-dimensional imagery, the deployment of multi-camera rigs or RGB-plus-depth sensors along the lines of Intel RealSense devices [38] could be harnessed to reconstruct a point cloud model and thereby pull out data tied to three-dimensional descriptors such as cross-sectional silhouette and petal architecture.

The mechanization of the workflow further opens the door to decentralizing trait appraisal, so that registering a new variety could one day resemble the electronic filing of a tax return. The capacity to reproduce outcomes from publicly accessible models would furnish a robust legal underpinning for a protocol of this nature.

© 2024 The Author(s). | Published by Agro Intelligence Research Press

Acknowledgements

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

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 20 Mar 2023 Revised: 25 May 2023 Accepted: 09 Jun 2023
Published online: 25 November 2024

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

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

2. Minervini M, Scharr H, Tsaftaris SA. Image analysis: The new bottleneck in plant phenotyping [applications corner]. IEEE Signal Process Mag. 2015;32(4):126-31. https://doi.org/10.1109/MSP.2015.2405111.

3. Fonteijn H, Afonso M, Mooij M, Faber N, Vroegop A, Polder G, et al. Automatic phenotyping of tomatoes in production greenhouses using robotics and computer vision: From theory to practice. Agronomy. 2021;11(8):1599.

4. Afonso M, Blok PM, Polder G, van der Wolf JM, Kamp J. Blackleg detection in potato plants using convolutional neural networks. IFAC-PapersOnLine. 2019;52(30):6-11. https://doi.org/10.1016/j.ifacol.2019.12.481.

5. Arad B, Balendonck J, Barth R, Ben-Shahar O, Edan Y, Hellström T, et al. Development of a sweet pepper harvesting robot. J Field Robot. 2020;37(6):1027-39.

6. Lehnert CF, English A, McCool C, Tow AW, Perez T. Autonomous sweet pepper harvesting for protected cropping systems. IEEE Robot Autom Lett. 2017;2(2):872-9.

7. Goodfellow I, Bengio Y, Courville A. Deep learning. MIT Press; 2016.

8. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436.

9. Wehrens R, Afonso M, Fonteijn H, Paulo J, Polder G, Zwinkels H, et al. Determining flower colors from images. Euphytica. 2023.
https://doi.org/10.1007/s10681-023-03258-2.

10. Afonso M, Mencarelli A, Polder G, Wehrens R, Lensink D, Faber N. Detection of tomato flowers from greenhouse images using colorspace transformations. In: Oliveira PM, Novais P, Reis LP, editors. EPIA 2019. Lecture Notes in Artificial Intelligence. Vol. 11804. Springer Nature Switzerland AG; 2019. p. 146-55.

11. Oppenheim D, Edan Y, Shani G. Detecting tomato flowers in greenhouses using computer vision. Int J Comput Electr Autom Control Inf Eng. 2017;11(1):104-9.

12. Dias PA, Tabb A, Medeiros H. Apple flower detection using deep convolutional networks. Comput Ind. 2018;99:17-28.
https://doi.org/10.1016/j.compind.2018.03.010.

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

14. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016:779-88.
https://doi.org/10.1109/CVPR.2016.91.

15. Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst. 2015:91-9.

16. Chen L, Papandreou G, Kokkinos I, Murphy K, Yuille AL. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. IEEE Trans Pattern Anal Mach Intell. 2018;40(4):834-48.
https://doi.org/10.1109/TPAMI.2017.2699184.

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

18. Hemming J. Current developments in greenhouse robotics and challenges for the future. Acta Hortic. 2020;1296:975-85.

19. Rath T, Kawollek M. Robotic harvesting of Gerbera Jamesonii based on detection and three-dimensional modeling of cut flower pedicels. Comput Electron Agric. 2009;66(1):85-92.
https://doi.org/10.1016/j.compag.2008.12.006.

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

21. Turgut K, Dutagaci H, Rousseau D. Rosesegnet: an attention-based deep learning architecture for organ segmentation of plants. Biosyst Eng. 2022;221:138-53.

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

© 2024 The Author(s). | Published by Agro Intelligence Research Press

23. Saitoh T, Kaneko T. Automatic recognition of wild flowers. In: Proceedings - International Conference on Pattern Recognition. 2000;2:507-10. https://doi.org/10.1109/ICPR.2000.906123.

24. Nilsback ME, Zisserman A. Automated flower classification over a large number of classes. In: 2008 Sixth Indian Conference on Computer Vision, Graphics and Image Processing. IEEE; 2008. p. 722-9.

25. Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M. U2-net: going deeper with nested U-structure for salient object detection. Pattern Recognit. 2020;106:107404.

26. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016:770-8.

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

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

29. Bradski G. The OpenCV library. Dr Dobb's J Softw Tools Prof Program. 2000;25(11):120-3.

30. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An Imperative Style, High-Performance Deep Learning Library. Adv Neural Inf Process Syst. 2019;32:8024-35.

31. Zhang L, Xu Z, Xu D, Ma J, Chen Y, Fu Z. Growth monitoring of greenhouse lettuce based on a convolutional neural network. Hortic Res. 2020;7(1):124.

32. Teimouri N, Dyrmann M, Nielsen PR, Mathiassen SK, Somerville GJ, Jørgensen RN. Weed growth stage estimator using deep convolutional neural networks. Sensors. 2018;18(5):1580. https://doi.org/10.3390/s18051580.

33. Tutz G. Ordinal regression: Areview and a taxonomy of models. WIREs Comput Stat. 2022;14(2):e1545. https://doi.org/10.1002/wics.1545.

34. Zheng M, Xu J, Shen Y, Tian C, Li J, Fei L, et al. Attention-based CNNs for image classification: A survey. J Phys Conf Ser. 2022;217(1):012068.

35. Lin Z, Zhang D, Tao Q, Shi D, Haffari G, Wu Q, et al. Medical visual question answering: A survey. Artif Intell Med. 2023;143:102611. https://doi.org/10.1016/j.artmed.2023.102611.

36. Hossain MDZ, Sohel F, Shiratuddin MF, Laga H. A comprehensive survey of deep learning for image captioning. ACM Comput Surv. 2019;51(6):1-36. https://doi.org/10.1145/3196375.

37. Wang K, Gou C, Duan Y, Lin Y, Zheng X, Wang FY. Generative adversarial networks: Introduction and outlook. IEEE/CAA J Autom Sin. 2017;4(4):588-98. https://doi.org/10.1109/JAS.2017.7510583.

38. Keselman L, Woodfill JI, Grunnet-Jepsen A, Bhowmik A. Intel RealSense stereoscopic depth cameras. Proc IEEE Conf Comput Vis Pattern Recognit Workshops. 2017:1-10. https://doi.org/10.1109/CVPRW.2017.167.

© 2024 The Author(s). | Published by Agro Intelligence Research Press