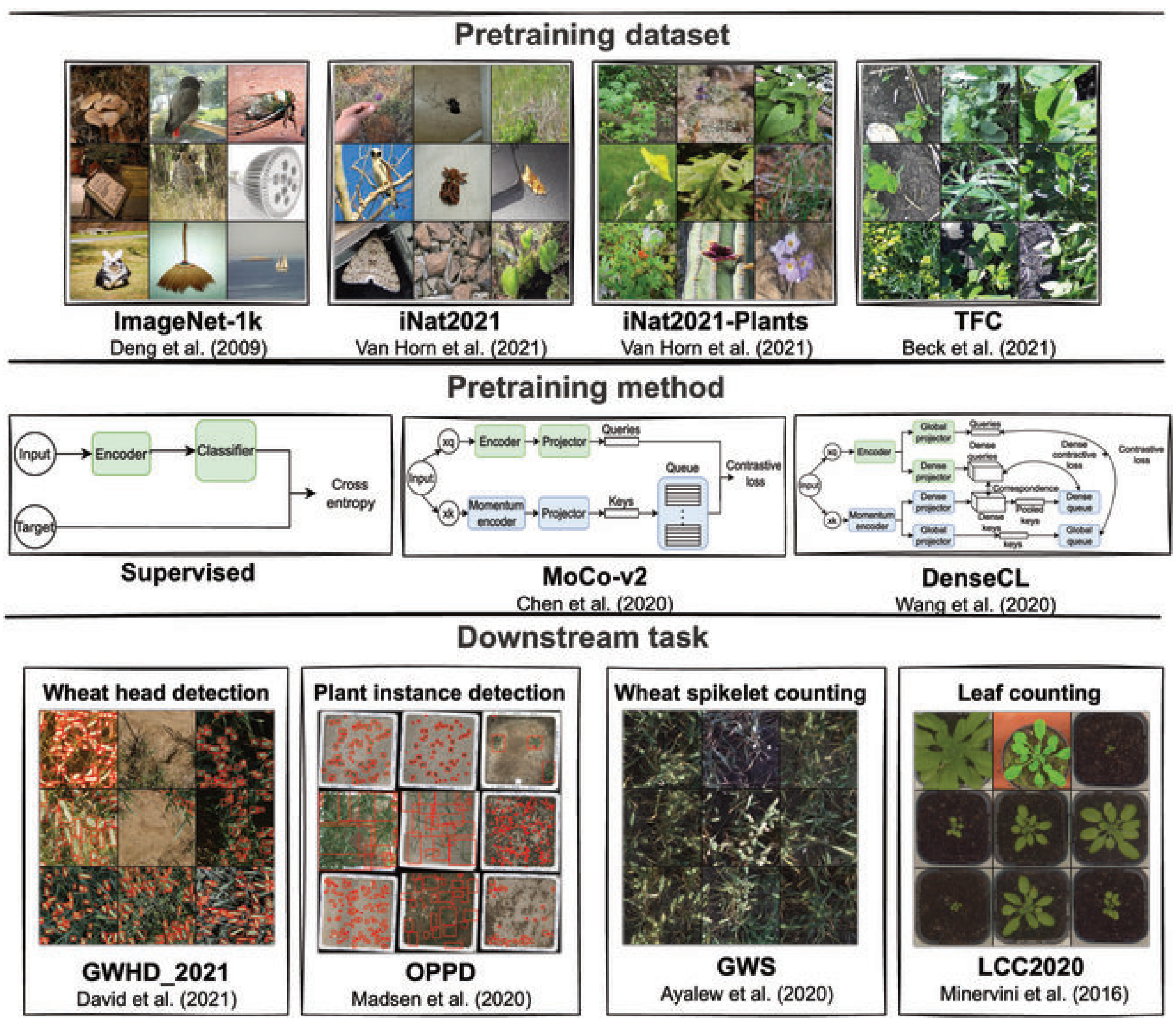
PHYS.ORG

Benchmarking deep-learning methods for more accurate plant-phenotyping

April 26 2023



(Top–bottom) The pretraining datasets with domains ranging from general-purpose concepts to field crop images, similar to those seen in the downstream tasks; the pretraining methods; and the downstream tasks, including object detection tasks (with bounding boxes shown in red) and counting tasks. We train an encoder with the pretraining datasets for each pretraining method and fine-tune the weights on each downstream task. Credit: *Plant Phenomics*

In crop-breeding, plant phenotyping is the detailed study of a plant's characteristic 'visible' or phenotypic features. It includes counting the number of plants generated by a crossing experiment and grading the features displayed by the offspring or progeny.

The progeny with the desirable traits is then crossed to produce the next generation of crops, and the process is repeated to enhance the crop variety. Conventional methods for plant phenotyping typically lack scalability, accuracy, and are immensely labor-intensive. This imposes a certain bottleneck on crop-breeding programs.

However, with technological advancements and the pertinent need for global food security to sustain a growing population, new methods are slowly taking the center stage. Among these are imaging-based techniques which capture photographs, extract features using machine learning tools, and compare the results with available databases, performing phenotyping tasks in a much shorter span of time and with higher accuracy.

Currently, most machine learning approaches adhere to the supervised learning framework with labeled datasets, which can be costly and time-consuming. The self-supervised learning (SSL) is a machine learning method that reduces the need for labeled data. Despite the surge in SSL research, there has been a dearth of SSL applications in image-based plant phenotyping tasks.

In a new study, a research team led by Associate Professor Ian Stavness from University of Saskatchewan, Canada, has benchmarked the performance of 2 SSL methods for improved plant phenotyping. The study, that claims that the self-supervised method may be more sensitive to pretraining dataset redundancy than the supervised ones, was published in *Plant Phenomics* .

Associate Prof. Stavness explains "These results highlight the importance of paying attention to dataset redundancy when training models for plant phenotyping tasks, especially when using SSL methods."

This study used wheat as a model crop to compare the conventional supervised (pre-training) methods to two SSL methods—momentum contrast (MoCo) v2 and dense contrastive learning (DenseCL). All learning methods were subject to 4 phenotyping tasks: wheat head detection, plant instance detection, wheat spikelet counting, and leaf counting. The team found that supervised pre-training produced the best-performing models for all the tasks except leaf-counting.

The contrastive SSL methods, unlike the supervised method, relied heavily on large labeled and annotated databases to perform phenotyping tasks. The algorithm is trained to pull positive samples closer and push negative samples apart, thereby increasing the strength of positive samples and training the algorithm to identify more of those samples.

While MoCo v2 functions on optimizing global image-level features of the sample, DenseCL focuses on local pixel-level features. Both methods showed comparable performance in the context of internal representations of training the models for the required tasks.

For the pre-training algorithms to perform well, a specialized but diverse dataset is essentially required. However, when faced with redundancy in large datasets, self-supervised methods perform better in terms of accuracy and sensitivity.

Given that the conclusions of the study are mostly drawn on empirical observations with few theoretical justifications, the authors want to extend their work into crop-breeding trials, which would involve more fine-grained information and assist them in training the algorithms for practical applications and large-scale commercialization.

Associate Prof. Stavness concludes by saying "SSL can potentially be used to learn a richer representation of plant phenotypes by aligning it with genotype and environment data in a joint embedding space. We hope that this benchmark or evaluation study will guide practitioners in developing better SSL methods for image-based plant phenotyping."

More information: Franklin C. Ogidi et al, Benchmarking Self-Supervised Contrastive Learning Methods for Image-Based Plant Phenotyping, *Plant Phenomics* (2023). DOI: 10.34133/plantphenomics.0037

Provided by NanJing Agricultural University

Citation: Benchmarking deep-learning methods for more accurate plant-phenotyping (2023, April 26) retrieved 20 June 2024 from https://phys.org/news/2023-04-benchmarking-deep-learning-methods-accurate-plant-phenotyping.html