Researchers use crowdsourcing to speed up data analysis in corn plants

August 27 2018

Participants in the study were shown images of corn plants and instructed to draw squares around the tassels. The images were used to create a computer algorithm capable of identifying tassels automatically. Credit: Iddo Friedberg

Teaching a machine to finish a complex task can save humans a lot of

time, effort and money. But first, the machine has to learn how, and that comes with plenty of its own challenges.

An interdisciplinary team of Iowa State University scientists turned to crowdsourcing, or relying on large groups of minimally trained people, to repeat a task often enough that researchers could formulate an algorithm that allows a computer to carry out that task automatically. In this case, the scientists wanted to teach a machine to identify the tassels of [corn plants](#) when given a vast number of photographic images to sort through.

The crowdsourcing effort produced results on par with those of trained plant scientists and resulted in an algorithm that will greatly reduce the time it takes to derive useful metrics from massive datasets. The researchers said the same approach could yield similar results for other crops. The academic journal *PLOS Computational Biology* recently published the team's results.

"This kind of automation could allow us to shave years off the time it would take to do this research otherwise," said Carolyn Lawrence-Dill, an associate professor of genetics, development and cell biology. "This approach can help us breed more effective crop varieties faster."

The research team used the Amazon Mechanical Turk, or MTurk, to find participants to take part in the study. MTurk is an online marketplace for [crowdsourcing](#), which allows employers to access a large pool of workers for simple tasks that nevertheless require human intelligence and can't be automated. Jonathan Kelly, an associate professor of psychology and research team member, said MTurk also provides a ready source of potential subjects for academic studies. The experiment also included student participants who were granted course credit.

All the study participants received instructions to identify tassels in dozens of images of corn by drawing a square around them. Baskar Ganapathysubramanian, professor of mechanical engineering, then used those labeled images to train a computer to identify tassels in similar corn images.

"You have to train a computer model to recognize the difference between, say, a leaf and a tassel," Ganapathysubramanian said. "One way to accomplish that is through machine learning, where you show instances of the object with a label over and over and the system builds its own rules to recognize that object on its own."

In the past, humans, often graduate students or someone with some expertise in a given discipline, would have to sift through the images one by one, a laborious process that can be sidestepped through machine learning, said Iddo Friedberg, associate professor of veterinary microbiology and preventive medicine and research team member. But Friedberg said the results obtained from MTurk proved to be as effective as those from trained plant experts. The results from the study's student participants were of slightly lesser quality but still good enough to work with the algorithm, he said. Friedberg called the results a testament to "the wisdom of the crowd."

A computer model capable of automatically identifying tassels or other plant parts can help plant scientists in a number of ways, Lawrence-Dill said. Scientists interested in how various weather conditions affect tasseling or how different corn hybrids perform can derive valuable insights more quickly with such a program. Perhaps even more importantly, she said, this method could perform a similar function for a range of traits for virtually any plant for which scientists possess a large stockpile of images.

The research was made possible after receiving funding from the

Presidential Initiative for Interdisciplinary Research. The ISU program encourages big thinking that ties together multiple disciplines in innovative ways. The research couldn't have happened without drawing on faculty members with expertise in statistics, bioinformatics, engineering, psychology and plant phenomics. But a team representing such diverse disciplines posed challenges as well. Some of the funding was also provided by the Plant Sciences Institute and an award from the National Science Foundation.

"Just explaining the terminology from our fields to one another took considerable effort," Lawrence-Dill said.

Provided by Iowa State University