Technique based on artificial intelligence permits automation of crop seed analysis

March 19 2021, by Janaína Simões

Credit: CC0 Public Domain

In Brazil, researchers affiliated with the Center for Nuclear Energy in Agriculture (CENA) and the Luiz de Queiroz College of Agriculture (ESALQ), both part of the University of São Paulo (USP), have

developed a methodology based on artificial intelligence to automate and streamline seed quality analysis, a process required by law and currently done manually by analysts accredited with the Ministry of Agriculture.

The group used light-based technology like that deployed in plant and cosmetics analysis to acquire images of the seeds. They then turned to machine learning to automate the image interpretation process, minimizing some of the difficulties of conventional methods. For example, for many species, optical imaging technology can be applied to an entire batch of seeds rather than just samples, as is the case currently. Furthermore, the technique is non-invasive and does not destroy the products analyzed or generate residues.

The light-based techniques consisted of chlorophyll fluorescence and multispectral imaging. Among plants that are relevant as both crops and experimental models, the researchers chose tomatoes and carrots produced in different countries and seasons and submitted to different storage conditions. They used seeds of the Gaucho and Tyna commercial tomato varieties produced in Brazil and the US, and seeds of the Brasilia and Francine carrot varieties produced in Brazil, Italy, and Chile.

The choice was based on the economic importance of these food crops, for which world demand is high and rising, and on the difficulties faced by growers in collecting their seeds. In both tomatoes and carrots, the ripening process is not uniform because the plants flower continuously and seed production is non-synchronous, so that seed lots may contain a mixture of immature and mature seeds. The presence of immature seeds is not easily detected by visual methods, and techniques based on machine vision can minimize this problem.

The researchers compared the results of their non-destructive analysis with those of traditional germination and vigor tests, which are destructive, time-consuming, and labor-intensive. In the germination test,

seed analysts separate samples, sow them to germinate in favorable temperature, water, and oxygen conditions, and verify the final quantity of normal seedlings produced in accordance with the rules established by the Ministry of Agriculture. Vigor tests are complementary and more sophisticated. The most common are based on the seed's response to stress and seedling growth parameters.

Besides the difficulties mentioned, traditional methods are time-consuming. In the case of tomatoes and carrots, for example, it can take up to two weeks to obtain results, which are also largely subjective, depending on the analyst's interpretation. "Our proposal is to automate the process as much as possible using chlorophyll fluorescence and multispectral imaging to analyze seed quality. This will avoid all the usual bottlenecks," said Clíssia Barboza da Silva, a researcher at CENA-USP and one of the authors of an article on the study published in *Frontiers in Plant Science*.

Silva is the principal investigator for the project supported by São Paulo Research Foundation - FAPESP. The lead author of the article is Patrícia Galletti, who conducted the study as part of her master's research and won the Best Poster Award in 2019 at the 7th Seed Congress of the Americas, where she presented partial results of the project.

Chlorophyll as a marker of quality

Chlorophyll is present in seeds, where it supplies energy for the storage of nutrients needed for development (lipids, proteins, and carbohydrates). Once it has fulfilled this function, the chlorophyll breaks down. "However, if the seed doesn't complete the maturation process, this chlorophyll remains inside it. The less residual chlorophyll, the more advanced the maturation process and the more and higher-quality the nutrients in the seed. If there's a lot of chlorophyll, the seed is immature

and its quality is poor," Silva said.

If light at a specific wavelength is shone on the chlorophyll in a seed, it does not transfer this energy to another molecule but instead re-emits the light at another wavelength, meaning that it fluoresces. This fluorescence can be measured, she explained. Red light can be used to excite chlorophyll and capture the fluorescence using a device that converts it into an electrical signal, producing an image comprising gray, black, and white pixels. The lighter areas correspond to higher levels of chlorophyll, indicating that the seed is immature and unlikely to germinate.

Artificial intelligence

In multispectral imaging, LEDs (light-emitting diodes) emit light in the visible portion of the spectrum as well as non-visible light (UV and near-infrared). To analyze seed quality based on reflectance, the researchers used 19 wavelengths and compared the results with quality assessment data obtained by traditional methods. The best results were obtained using near-infrared in the case of carrot seeds and UV in the case of tomato seeds.

Seeds contain proteins, lipids and sugars that absorb part of the light emitted by the LEDs and reflect the rest. The reflected light is captured by a multispectral camera, and the image captured is processed to separate the seeds from the support in the device, which corresponds to black pixels with zero value, while the seeds are gray-scale. The values of the pixels in the image of a seed correspond to its chemical composition.

"We don't work with an average result for a sample. We perform individualized extraction for each seed," Silva said. "The larger the amount of a given nutrient the seed contains, the more light of a specific wavelength it absorbs so that less is reflected. A seed with a smaller

nutrient content contains fewer light-absorbing molecules. This means its reflectance is higher, although this varies according to its components, which behave differently depending on the light wavelength used."

An algorithm identifies the wavelength that obtains the best result. The process provides information about the seed's chemical composition, from which its quality can be inferred.

For the researchers, it was not enough to reach the imaging stage, as this is still an operation that requires human observation. "We then deployed chemometrics, a set of statistical and mathematical methods used to classify materials chemically," Silva said. "The idea was that the equipment should classify quality on the basis of the image it captured." The methods used by the scientists in this study are widely used in medicine and the food industry.

Next, they leveraged machine learning to test the models created using chemometrics. "We taught the model to identify high-quality and low-quality seeds. We used 70% of our data to train the model, and used the remaining 30% for validation," Silva said. Quality classification accuracy ranged from 86% to 95% in the case of tomato seeds, and from 88% to 97% in the case of carrot seeds.

The two main techniques were both accurate and time-saving, given the speed of image capture. The chlorophyll fluorescence instrument captured one image per second, while the multispectral imaging analyzer processed 19 images in five seconds.

Unexpected results

An unexpected result produced in the course of the project proved highly important. Chlorophyll fluorescence and multispectral imaging are also efficient techniques for plant variety screening, an essential part

of seed lot evaluation to avoid economic losses. "Growers buy seeds with the expectation of a certain crop yield, but production will be affected if seeds with different genetic characteristics aren't properly separated," Silva said.

Screening is currently done by analysts trained in the skills needed to grade seeds by color, shape, and size, as well as molecular markers where possible. In the study, both techniques proved efficient to separate carrot varieties but multispectral imaging was unsatisfactory in the case of tomato varieties.

"The study produced novel results with regard to the use of fluorescence to screen varieties," Silva said. "We found no prior research in which fluorescence was used for this purpose. Some studies show multispectral imaging to be efficient for this purpose, but not with the instrument we used."

Instrument sharing

A good way to transfer the knowledge produced by the research to the productive sector, Silva said, would be to have firms develop the equipment for sale to seed producers. "It would be possible to use the results of our research to develop an instrument that used only UV light to characterize tomato seed quality and bring it to market, for example," she surmised.