AI reveals hidden traits about our planet's flora to help save species

June 20 2023, by Lilly Matson
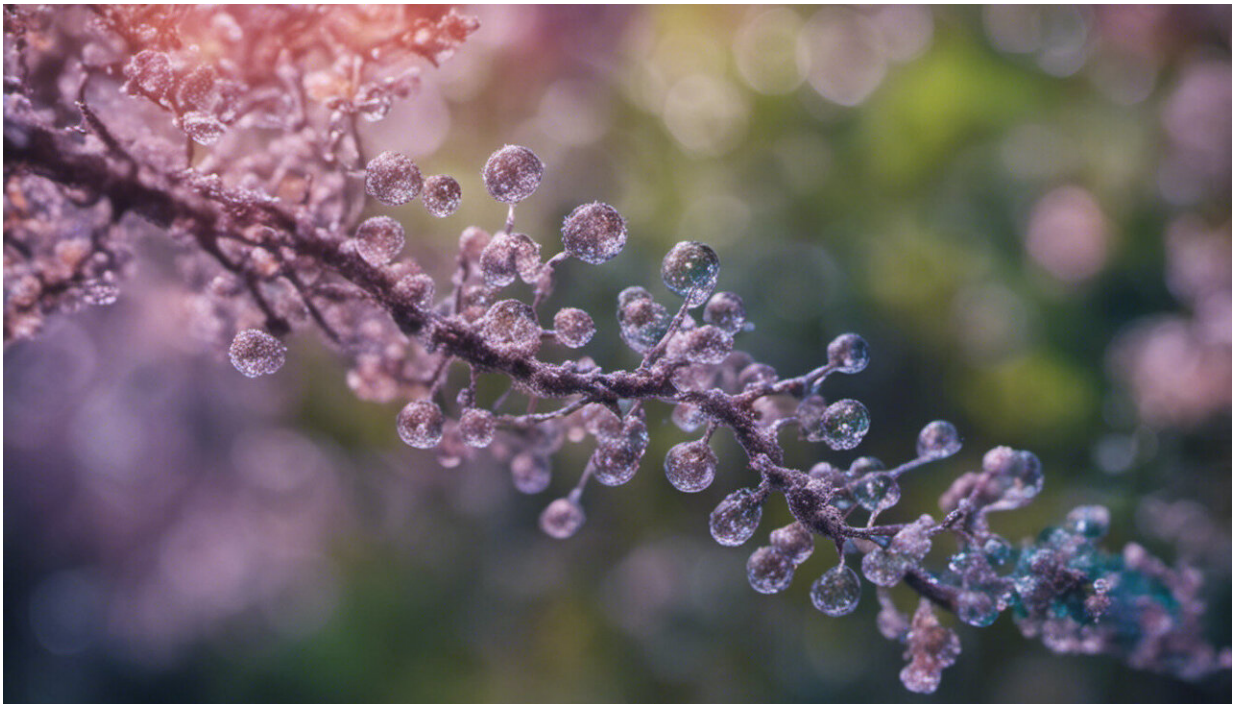
Scientists from UNSW and Botanic Gardens of Sydney have trained AI to unlock data from millions of plant specimens kept in herbaria around the world in order to study and combat the impacts of climate change on flora.

"Herbarium collections are amazing time capsules of plant specimens," says lead author on the study, Associate Professor Will Cornwell. "Each year over 8,000 specimens are added to the National Herbarium of New South Wales alone, so it's not possible to go through things manually anymore."

Using a new machine learning algorithm to process more than 3,000 [leaf samples](), the team discovered that contrary to frequently observed interspecies patterns, [leaf size]() doesn't increase in [warmer climates]() within a [single species]().

Published in the *American Journal of Botany*, this research not only reveals that factors other than climate have a strong effect on leaf size within a plant species, but demonstrates how AI can be used to transform static specimen collections and to quickly and effectively document climate change effects.

Herbarium collections move to the digital world

Herbaria are scientific libraries of plant specimens that have existed since at least the 16th century. "Historically, a valuable scientific effort was to go out, collect plants, and then keep them in a herbarium. Every record has a time and a place and a collector and a putative species ID," says A/Prof. Cornwell, a researcher at the School of BEES and a member of UNSW Data Science Hub.

A couple of years ago, to help facilitate scientific collaboration, there was a movement to transfer these collections online.

"The herbarium collections were locked in small boxes in particular places, but the world is very digital now. So to get the information about all of the incredible specimens to the scientists who are now scattered across the world, there was an effort to scan the specimens to produce

high resolution digital copies of them."

The largest herbarium imaging project was undertaken at the Botanic Gardens of Sydney when over 1 million plant specimens at the National Herbarium of New South Wales were transformed into high-resolution digital images.

"The digitization project took over two years and shortly after completion, one of the researchers—Dr. Jason Bragg—contacted me from the Botanic Gardens of Sydney. He wanted to see how we could incorporate machine learning with some of these high-resolution digital images of the Herbarium specimens."

Dr. Bragg says, "I was excited to work with A/Prof. Cornwell in developing models to detect leaves in the plant images, and to then use those big datasets to study relationships between leaf size and climate."
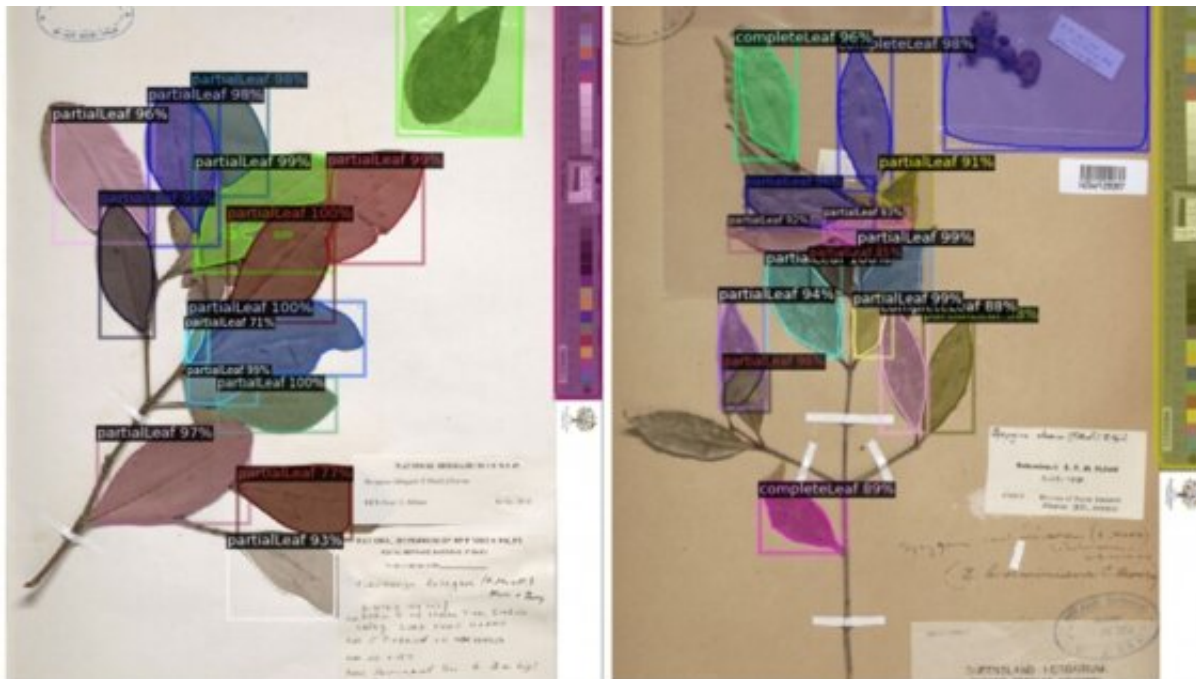
The machine learning algorithm developed by the research team measures and

identifies plant specimens. Credit: University of New South Wales

'Computer vision' measures leaf sizes

Together with Dr. Bragg at the Botanic Gardens of Sydney and UNSW Honors student Brendan Wilde, A/Prof. Cornwell created an algorithm that could be automated to detect and measure the size of leaves of scanned herbarium samples for two plant genera—Syzygium (generally known as lillipillies, brush cherries or satinas) and Ficus (a genus of about 850 species of woody trees, shrubs and vines).

"This type of AI is called a convolutional neural network, also known as Computer Vision," says Cornwell. The process essentially teaches the AI to see and identify the components of a plant in the same way a human would.

"We had to build a [training data](#) set to teach the computer: this is a leaf, this is a stem, this is a flower," says Cornwell. "So we basically taught the computer to locate the leaves and then measure the size of them.

"Measuring the size of leaves is not novel, because lots of people have done this. But the speed with which these specimens can be processed and their individual characteristics can be logged is a new development."

A break in frequently observed patterns

A general rule of thumb in the botanical world is that in wetter climates, like tropical rainforests, the leaves of plants are bigger compared to drier climates, such as deserts.

"And that's a very consistent pattern that we see in leaves between

species all across the globe," says Cornwell. "The first test we did was to see if we could reconstruct that relationship from the machine-learned data, which we could. But the second question was, because we now have so much more data than we had before, do we see the same thing within species?"

The machine learning algorithm was developed, validated, and applied to analyze the relationship between leaf size and climate within and among species for Syzygium and Ficus plants.

The results from this test were surprising—the team discovered that while this pattern can be seen between different plant species, the same correlation isn't seen within a single species across the globe, likely because a different process, known as gene flow, is operating within species. That process weakens plant adaptation on a local scale and could be preventing the leaf size-climate relationship from developing within species.

Using AI to predict future climate change responses

The machine learning approach used here to detect and measure leaves, though not pixel perfect, provided levels of accuracy suitable for examining links between leaf traits and climate.

"But because the world is changing quite fast, and there is so much data, these kinds of machine learning methods can be used to effectively document climate change effects," says Cornwell.

What's more, the machine learning algorithms can be trained to identify trends that might not be immediately obvious to human researchers. This could lead to new insights into plant evolution and adaptations, as well as predictions about how plants might respond to future effects of climate change.

More information: Brendan C. Wilde et al, Analyzing trait-climate relationships within and among taxa using machine learning and herbarium specimens, *American Journal of Botany* (2023). DOI: 10.1002/ajb2.16167

Provided by University of New South Wales