Multidimensional image processing and analysis in R

June 2 2014, by Linda Vu

Talita Perciano

Many researchers believe that an esoteric, open-source programming language for statistical analysis—called R—could pave the way for open science. Today, thousands of international scientists are participating in the R development community—contributing new tools and libraries, including some that branch away from statistical analysis. And that number is rapidly growing.

One such contributor is Talita Perciano, a postdoctoral researcher in the Lawrence Berkeley National Laboratory's (Berkeley Lab's) Computational Research Division's Visualization Group. As a graduate student, Perciano contributed one of the first image-processing

tools—called R Image Processing Analysis (RIPA)—to the community. Now with big science datasets in mind, she's updated the existing tool with improved features for complex data analysis.

"When I started working on RIPA in 2005 as an undergraduate student, the idea was to create an image processing package to analyze satellite images. My advisor and I developed the tools, which helped us to write several manuals and lectures on the topic. This material was the basis for our book Introduction to Image Processing using R," says Perciano. "At Berkeley Lab, we are dealing with bigger and more complex experimental and simulated scientific datasets. So I've been updating RIPA to run on massively parallel clusters at the National Energy Research Scientific Computing (NERSC) facility and perform much more advanced image processing tasks."

Perciano notes that the most recent RIPA release allows its users to do some essential image processing in parallel, like thresholding, changing contrast and brightness, and filtering. But the next version (to be released later this year) will be able to perform more complex image processing tasks in parallel. It will include new algorithms for pattern recognition and feature extraction, and it will be able to handle three-dimensional images. Perciano is also working on a new book about how to use the updated tools

Traditionally, privately owned tools like MATLAB (the mathematical computing software) and SAS (the statistical tool) have been necessities in research laboratories, similar to the way Microsoft Office is in office settings. But these tools can be expensive, and they have some limitations.

In fact, this is how Berkeley Lab Data Scientist Daniela Ushizima discovered RIPA. "I was working with a NERSC user on a project to visualize graphical patterns from massive datasets, such as Flickr images

and TIME magazine covers for a cultural analytics project," she says. "The number of images was large and required parallel processing. Because of the scale of the problem, MATLAB was not an option, so we had to look for other analytics tools."

After searching the literature for alternatives, Ushizima found RIPA. "This tool was perfect because R is the lingua franca for statistical analysis and RIPA gave us many image processing capabilities inside R," says Ushizima. "Because R is open source, there is an extensive community of users and developers to support the creation of R-based algorithms and packages."

Today, R is used in a range of scientific disciplines from astronomy to genomics, and even in drug development. Because it is an open-source statistical framework, it allows users to quickly share techniques with other R users, as well as reproduce and reuse the techniques they have discovered. New codes and techniques are shared through groups like the Comprehensive R Archive Network (CRAN), which is where Perciano published her package.

"RIPA is one of the best image processing/analysis packages in R," says Ushizima, who works with Perciano in image analysis and recognition at LBNL.

 More information: For more information about R:

www.fastcolabs.com/3028381/how … en-source-to-science
www.revolutionanalytics.com/what-r
www.r-project.org/

More information about RIPA:

cran.r-project.org/web/packages/ripa/index.html

Provided by Lawrence Berkeley National Laboratory